

**HELISSON FAORO**

**Análise genética e mutagênese sítio dirigida do gene *glnE* de  
*Herbaspirillum seropedicae***

Projeto de Monografia apresentado à Coordenação do curso de Ciências Biológicas, Setor de Ciências Biológicas, Universidade Federal do Paraná, como parte integrante para a obtenção da modalidade de Bacharelado em Ciências Biológicas.

Orientadores:

Prof<sup>a</sup> Dra Maria Berenice Reynaud Steffens

Prof. Dr Fábio de Oliveira Pedrosa

Prof<sup>a</sup> Dra Liu Un Rigo

Prof<sup>a</sup> Dra Roseli Wassem

**CURITIBA**

**2004**

UNIVERSIDADE FEDERAL DO PARANÁ  
SETOR DE CIÊNCIAS BIOLÓGICAS  
DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR

**TÍTULO DO PROJETO DE MONOGRAFIA**

Análise genética e mutagênese sítio dirigida do gene *glnE* de *Herbaspirillum seropedicae*.

**ALUNO**

HELISSON FAORO

**ORIENTADORES**

Prof<sup>ª</sup>. Dra. MARIA BERENICE REYNAUD STEFFENS

Prof. Dr FÁBIO DE OLIVEIRA PEDROSA

Prof<sup>ª</sup>. Dra. LIU UN RIGO

Departamento de Bioquímica e Biologia Molecular

Prof<sup>ª</sup>. Dra ROSELI WASSEM

Departamento de Genética

**COLABORADORES:**

Prof. Dr.Emanuel Maltempi de Souza

Prof<sup>ª</sup>. Dra. Leda Satie Chubatsu

Profa. Dra Elaine Machado Benelli

Prof. Dr. Leonardo Magalhães Cruz

## RESUMO

GlnE é uma enzima efetora bifuncional que atua regulando a cascata do metabolismo de nitrogênio controlando a atividade da enzima Glutamina Sintetase (GS). GS é uma das enzimas centrais envolvidas na assimilação de nitrogênio catalisando a conversão de glutamato e amônio em glutamina. A proteína GlnE regula a atividade enzimática de GS através da adenililação/desadenilidação de suas subunidades. Em *H. seropedicae*, uma bactéria diazotrófica endofítica, o gene *glnE* foi identificado durante o programa de sequenciamento genômico (GENOPAR). Nesse organismo, o gene *glnE* codifica para uma proteína de 927 aminoácidos que apresenta alta similaridade com o gene *glnE* de outros organismos. O gene *glnE* é, aparentemente, monocistônico e constitutivamente expresso. Dois domínios homólogos conservados foram identificados na proteína GlnE: um na região C-terminal que possui o sítio de desadenililação e outro na região C-terminal que possui o sítio de adenililação. Um cassete de canamicina foi inserido no gene *glnE* clonado no vetor pUC18. O plasmídeo resultante (pHF16), foi eletroporado na estirpe selvagem SMR1 de *H. seropedicae* gerando duas estirpes mutantes por inserção cromossomal (HEL1 e HEL2). O mutante HEL1 apresentou atividade nitrogenase reduzida em relação à estirpe selvagem SMR1.

## ÍNDICE

<b>1. INTRODUÇÃO .....</b>	<b>1</b>
1.1 Fixação biológica de nitrogênio.....	1
1.2 Metabolismo de nitrogênio.....	2
1.3 Sistema <i>Ntr</i> .....	4
1.4 Regulação da atividade da GS .....	6
1.5 Estrutura da proteína GlnE em <i>E. coli</i> .....	8
1.6 <i>Herbaspirillum seropedicae</i> .....	10
<b>2. OBJETIVOS.....</b>	<b>13</b>
2.1 Objetivos gerais .....	13
2.2 Objetivos específicos.....	13
<b>3. MATERIAL E MÉTODOS.....</b>	<b>14</b>
3.1 Microrganismos e plasmídeos.....	14
3.2 Condições e meios de cultivo das bactérias .....	14
3.2.1 Antibióticos .....	16
3.3 Manipulação de DNA .....	16
3.3.1 Purificação de DNA plasmidial.....	16
3.3.2 Purificação de DNA genômico de <i>H. seropedicae</i> .....	17
3.3.3 Sequenciamento de DNA plasmidial .....	17
3.4 Análise da sequência do gene <i>glnE</i> de <i>Herbaspirillum seropedicae</i> .....	18
3.5 Preparo de células eletrocompetentes .....	19
3.6 Transformação bacteriana.....	19
3.7 Seleção de mutantes cromossômicos <i>glnE</i> de <i>Herbaspirillum seropedicae</i> .....	19
3.7.1 Isolamento do cassete de canamicina .....	20
3.8 Mutagenese sítio dirigida.....	20
3.9 Hibridização do DNA .....	20
3.10 Determinação da atividade da nitrogenase.....	21



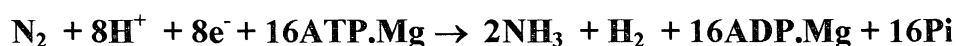
<b>4 RESULTADOS E DISCUSSÃO .....</b>	<b>1</b>
4.1 Identificação de clones no banco de dados do Programa GENOPAR contendo o gene <i>glnE</i> de <i>Herbaspirillum seropedicae</i> .....	23
4.2 Análise de um fragmento de DNA cromossomal de <i>H. seropedicae</i> contendo o gene <i>glnE</i> .....	28
4.2.1 Análise do gene <i>glnE</i> de <i>H. seropedicae</i> .....	31
4.2.2 Análise da proteína GlnE .....	31
4.3 Mutagênese do gene <i>glnE</i> .....	42
4.4 Caracterização fisiológica dos mutantes.....	54
<b>5. CONCLUSÃO.....</b>	<b>56</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>57</b>
<b>ANEXOS .....</b>	<b>64</b>

## 1. INTRODUÇÃO

### 1.1 Fixação biológica de nitrogênio

Fixação Biológica de Nitrogênio é a capacidade que diferentes grupos de procariotos tem de reduzir o nitrogênio atmosférico a amônio; numa reação catalisada pelo complexo enzimático da nitrogenase.

A reação global da redução do  $N_2$  pela nitrogenase é a seguinte:



Os microrganismos que apresentam esta propriedade são denominados diazotrofos (POSTGATE et al., 1982) e contribuem com aproximadamente 60% do nitrogênio fixado no planeta, sendo, portanto essenciais para a manutenção do ciclo do nitrogênio (BURNS & HARDY et al., 1975).

A ativação da fixação de nitrogênio pelos organismos diazotróficos requer condições ambientais favoráveis. Dentre estas, incluem-se ausência de nitrogênio fixado, baixas tensões de oxigênio, presença de molibdênio e temperatura adequada para cada organismo (POSTGATE et al., 1982). Tanto o excesso de nitrogênio fixado quanto de oxigênio reprimem a expressão dos genes da fixação de nitrogênio nos diversos organismos diazotróficos estudados (EADY et al., 1986; POSTGATE et al., 1982).

O nitrogênio está diretamente relacionado com a produtividade agrícola (MALIK et al., 1985). Porém, na maioria dos solos, a disponibilidade deste nutriente é pequena quando comparada com a necessidade da cobertura vegetal (EPSTEIN et al., 1975) e por este motivo, deve ser constantemente repostado através do uso de fertilizante, ou pela otimização do processo de fixação biológica de nitrogênio (POSTGATE et al., 1982).

## 1.2 Metabolismo de nitrogênio

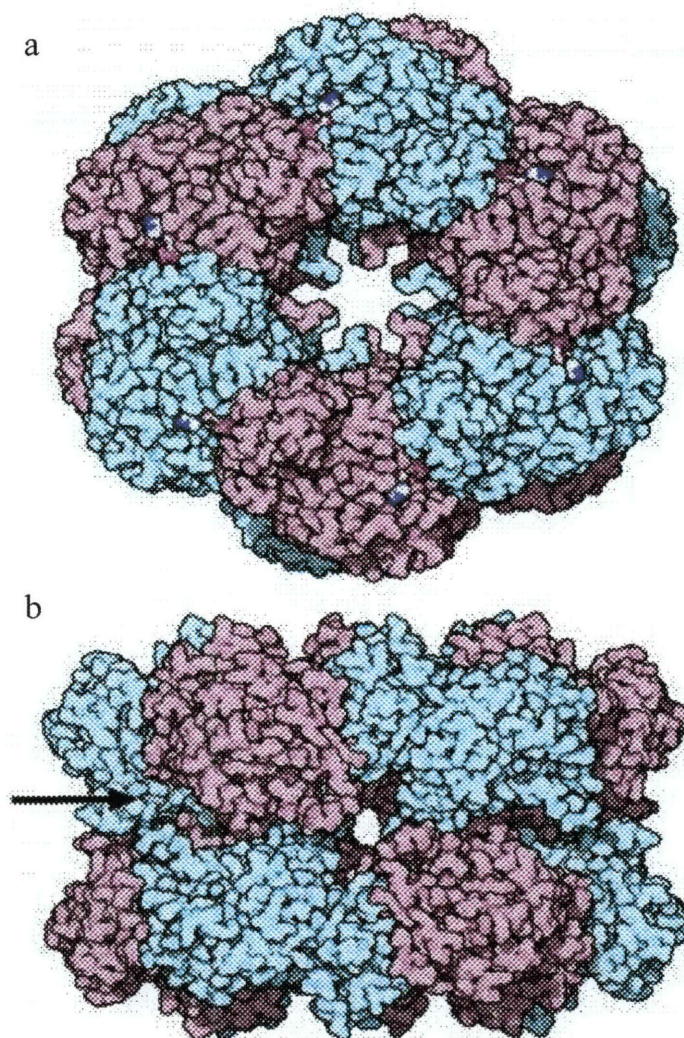
Por ser necessário para a produção de aminoácidos, nucleotídeos, NAD, lipopolisacarídeos e peptidioglicanos, o nitrogênio é um dos muitos elementos essenciais para a vida. Dentre os seres vivos, as bactérias desenvolveram mecanismos para a obtenção de nitrogênio a partir de várias fontes, desde amônio até dinitrogênio atmosférico ( $N_2$ ) (ARCONDÉGUY et al., 2001). Amônio é quase sempre preferido como fonte de nitrogênio, podendo ser assimilado diretamente como glutamato e glutamina. Fontes orgânicas de nitrogênio, como os aminoácidos, precisam sofrer desaminação, e fontes inorgânicas, como  $NO_3^-$ ,  $NO_2^-$  e  $N_2$ , precisam ser reduzidas antes da assimilação (REITZER, 1996).

Em virtualmente todas as células, glutamato e glutamina servem como o doador chave de nitrogênio para reações de biossíntese. Há dois caminhos principais que facilitam a incorporação do nitrogênio presente no glutamato e na glutamina. Um desses caminhos é mediado pela enzima glutamato desidrogenase (GDH). O outro, e mais sensível, é mediado pelas proteínas glutamina sintetase (GS) e glutamato sintase (GOGAT). A glutamina sintetase converte glutamato e amônia em glutamina, e a glutamato sintase transfere um grupo amido da glutamina para o  $\alpha$ -cetoglutarato produzindo duas moléculas de glutamato. A diferença de sensibilidade entre as duas vias se deve ao fato de que a enzima GDH apresenta um alto  $K_m$  para amônia e, desse modo, atua apenas em altas concentrações desse substrato (MERRICK & EDWARDS, 1995).

A GS, produto do gene *glnA*, se apresenta comumente como uma enzima dodecamérica com subunidades idênticas de aproximadamente 55 kDa cada (figura 1). A estrutura cristalizada dessa enzima revelou que as doze subunidades idênticas se arranjam em dois anéis sobrepostos unidos por pontes de hidrogênio e interações hidrofóbicas (YAMASHITA et al, 1989). A enzima GOGAT por sua vez é formada por duas subunidades não idênticas que, em *Escherichia coli*, possuem

massa molecular de 53 e 135 kDa e são codificadas pelos genes *glbD*. (OLIVER et al., 1987).

**FIGURA 1: Estrutura da enzima Glutamina Sintetase.**



Glutamina Sintetase em vista superior (a) e vista lateral (b). As subunidades da enzima estão destacadas em verde e rosa.

### 1.3 Sistema *Ntr*

A combinação de experimentos bioquímicos e genéticos levaram ao reconhecimento, em enterobactérias, de que a expressão de proteínas relacionadas ao catabolismo e à assimilação de nitrogênio é coordenada por um sistema regulatório central (MERRICK & EDWARDS, 1995). Esse sistema global de regulação de nitrogênio (sistema *Ntr*) foi considerado inicialmente como sendo constituído de quatro proteínas: GlnD (UTase/UR), GlnB (proteína PII) e proteínas de dois componentes histidina quinase NtrB e NtrC, o qual, em enterobactérias, é co-transcrito no operon *glnA-ntrBC*. Posteriormente foi descoberto que GlnK, em *E. coli* e muitas outras bactérias, também participa do sistema (VAN HEESWIJK et al., 1996). Atualmente o sistema *Ntr* engloba as proteínas GlnB (PII), NtrB, NtrC, GS (glutamina sintetase), GlnD (uridilil transferase e removedora de uridilil), GlnE (adenilil transferase e removedora de adenilil) e GlnK. O sistema Ntr de enterobactérias é usado como modelo por ter sido intensamente estudado.

O controle da atividade de NtrC em resposta aos níveis de nitrogênio é mediada por NtrB. Neste sistema o produto do gene *ntrB* (NtrB) é uma proteína sensora dos níveis intracelulares de amônia e controla a atividade do produto do gene *ntrC* (NtrC) através de reações de fosforilação/desfosforilação (REITZER & MAGASANIK, 1983). A proteína NtrC por sua vez, é responsável pela regulação da transcrição de uma série de genes envolvidos na assimilação de nitrogênio. (MERRICK, 1983)

Em baixas concentrações de amônia a proteína UTase/UR converte PII para a forma uridililada PII-UMP através da ligação de um grupo UMP ao resíduo tirosina 51 da proteína PII (SON & RHEE, 1987). PII-UMP não interage com a proteína NtrB que, quando livre, sofre autofosforilação no resíduo histidina 139 (WEISS e MAGASANIK, 1988). Este grupamento fosforil é então transferido da proteína NtrB para o resíduo Asp54 do domínio N-terminal da proteína NtrC, que somente na forma NtrC-P atua como ativador de transcrição (WEISS et al., 1991).



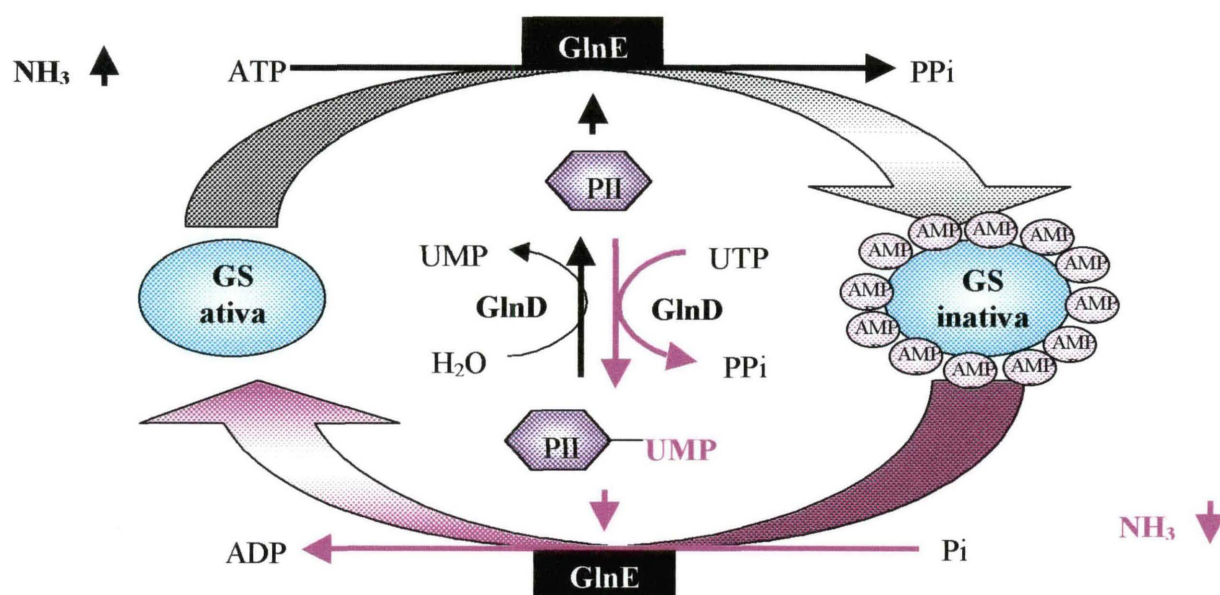


### 1.4 Regulação da atividade da GS

Em *E. coli* e bactérias relacionadas foram descritos três mecanismos distintos de regulação da GS. Primeiro, a concentração intracelular da enzima é regulada em resposta ao nível intracelular de nitrogênio. Essa resposta é necessária para o início da transcrição do gene *glnA*, já que o promotor *glnAp2* desse gene é regulado por nitrogênio (REITZER & MAGASANIK, 1985). Em segundo lugar, a atividade da enzima é regulada por modificação covalente reversível (STADTMAN et al., 1980) através da adição de grupos adenilil em cada subunidade da enzima, sendo a forma GS-AMP inativa. Finalmente, a atividade da GS é controlada alostéricamente por uma retroalimentação inibitória acumulativa por oito moléculas: triptofano, histidina, carbamil fosfato, glucosamina-6-fosfato, CTP, AMP, alanina e glicina (NINFA et al., 1995).

A regulação da atividade enzimática da GS envolve principalmente duas enzimas bifuncionais, uridilil transferase (GlnD ou UTase) e adenilil transferase (GlnE ou ATase). Quando o nível de nitrogênio é baixo, GlnD (enzima sensora) converte PII em PII-UMP por ligação covalente de um UMP (proveniente de um UTP), no resíduo de tirosina na posição 51 (SON & RHEE, 1987). PII-UMP interage com GlnE, produto do gene *glnE*, estimulando sua atividade de desadenilação (ANDERSON & STADTMAN, 1970). Em condições de excesso de nitrogênio, a atividade de remoção do grupo uridilil é estimulada, e PII-UMP é convertida em PII. A proteína PII interage com GlnE, estimulando a atividade de adenililação que converte a GS em GS-AMP inativa (figura 3), consequência da ligação de um grupo adenilil a cada subunidade da GS (MAGASANIK, 1993).

**FIGURA 3: Regulação da atividade de GS**



Modelo geral da regulação de GS por GlnD e GlnE. As setas pretas indicam uma condição de excesso de nitrogênio e as setas vermelhas indicam uma condição de nitrogênio escasso.



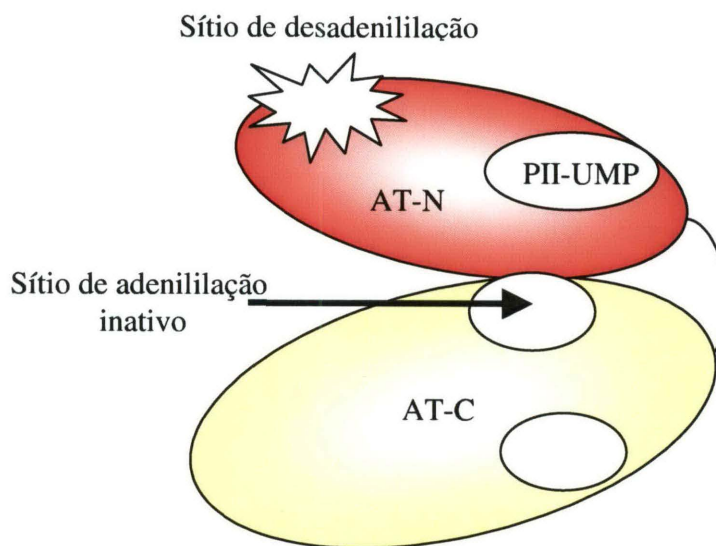
### 1.5 Estrutura da proteína GlnE em *E. coli*

A proteína GlnE é codificada pelo gene *glnE* localizado a 68,9 minutos no genoma da *E. coli* e é constitutivamente expressa (MUSE & BENDER, 1992). É uma enzima bifuncional que promove tanto a adenililação quanto a desadenililação da GS inativando-a ou ativando-a respectivamente.

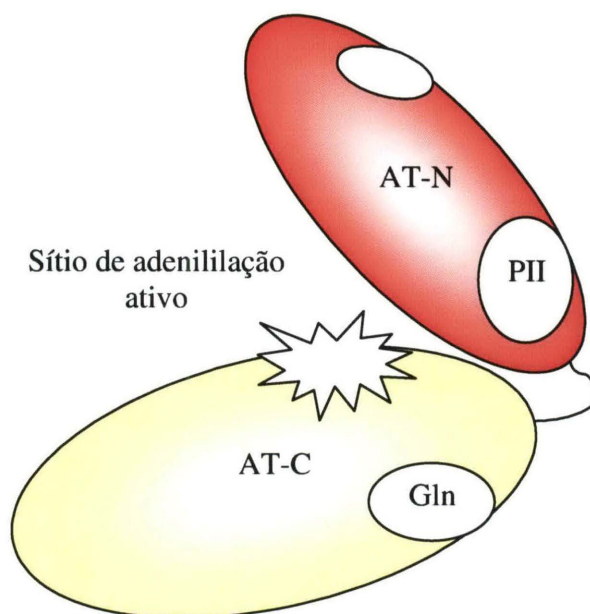
O gene *glnE* codifica um polipeptídio de cadeia simples com massa molecular de aproximadamente 115 kDa (CABAN & GISBURG, 1976). A determinação de sua sequência de nucleotídeos e a análise de seu produto indicam que a região N-terminal de GlnE possui alta homologia com a sua região C-terminal (van HEESWIJK et al., 1993). JAGGI e colaboradores (1997) produziram dois polipeptídios, AT-N e AT-C, correspondendo às regiões N-terminal e C-terminal de GlnE respectivamente. Neste trabalho foi demonstrado que ambos os domínios, AT-N e AT-C, possuem atividade catalítica, sendo que o primeiro é responsável pela desadenililação e o segundo pela adenililação da GS. Foi então proposto um modelo para as atividades de adenililação/desadenililação (JAGGI et al., 1997). O sítio que promove a desadenililação, localizado no domínio AT-N, possui também um sítio para ligação de PII e PII-UMP, enquanto que o sítio que realiza a adenililação, localizado no domínio AT-C, possui um sítio para ligação da glutamina. A ligação conjunta de PII-UMP com  $\alpha$ -cetoglutarato e ATP no domínio AT-N é necessária para o início da atividade de desadenilação. Com a ligação de PII-UMP, GlnE assume uma conformação “fechada” que, junto com a atividade de desadenilação, bloqueia o sítio ativo de adenililação no domínio AT-C. Entretanto, durante a atividade de adenililação, a ligação de PII faz com que GlnE assumira uma conformação “aberta” expondo o sítio ativo responsável pela adenililação, a qual é ativada pela ligação de glutamina no domínio AT-C (figura 4).

**FIGURA 4: Regulação da atividade da enzima GlnE.**

**a) Baixa concentração de amônia**



**b) Alta concentração de amônia**



Modelo proposto por JAGGI (1997). a) altas concentrações de amônia e b) baixas concentrações de amônia. AT-N representa o domínio N-terminal e AT-C o domínio C-terminal da proteína GlnE. Estão indicados também os sítios para ligação de glutamina (Gln) e para PII/PII-UMP.

### 1.6 *Herbaspirillum seropedicae*

*Herbaspirillum seropedicae*, originalmente isolado de superfícies desinfetadas de raízes e colmos de milho, arroz e sorgo, foi inicialmente identificado como uma nova espécie do gênero *Azospirillum* (BALDANI et al., 1984). Posteriormente, foram evidenciadas diferenças em relação a dimensão das células, número de flagelos e coloração das colônias (BALDANI et al., 1986). FALK e colaboradores (1986), com base em experimentos de hibridização rRNA/DNA mostraram a pequena relação filogenética entre *H. seropedicae*, *Azospirillum* sp e *Aquaspirillum itersonii*; e BALDANI e colaboradores (1986) demonstram que a homologia rRNA/DNA entre *H. seropedicae* e *Azospirillum* spp era menor que 24%. Estes últimos propuseram então a criação do gênero *Herbaspirillum*. Diferenças entre *Herbaspirillum* e *Azospirillum* também foram confirmadas por seus padrões protéicos de membrana em SDS-PAGE (DIANESE et al., 1989).

*H. seropedicae* é classificado como um organismo diazotrófico gram-negativo, pertencente a subclasse  $\beta$  das proteobactérias. As células são vibrióides, podendo variar algumas vezes para helicoidais com tamanho que varia de 1,5 a 5  $\mu$ m, conforme o meio de cultivo utilizado. Possui de um a três flagelos bipolares e seu movimento pode ser observado por microscopia óptica a fresco. *H. seropedicae* apresenta um metabolismo respiratório típico onde o açúcar não é fermentado. Fixa nitrogênio atmosférico sob condições microaeróbicas e cresce com dinitrogênio gasoso como única fonte de nitrogênio (BALDANI et al., 1986). Ácidos orgânicos são as principais fontes de carbono, mas glucose, galactose e arabinose também podem ser utilizadas. Sacarose, entretanto, não é oxidada. O pH ótimo de crescimento é bastante amplo, variando desde 5,3 até 8,0. A temperatura ótima de crescimento fica em torno de 35°C (BALDANI et al., 1986; FU E BURRIS 1989; KLASSEN et al., 1997).

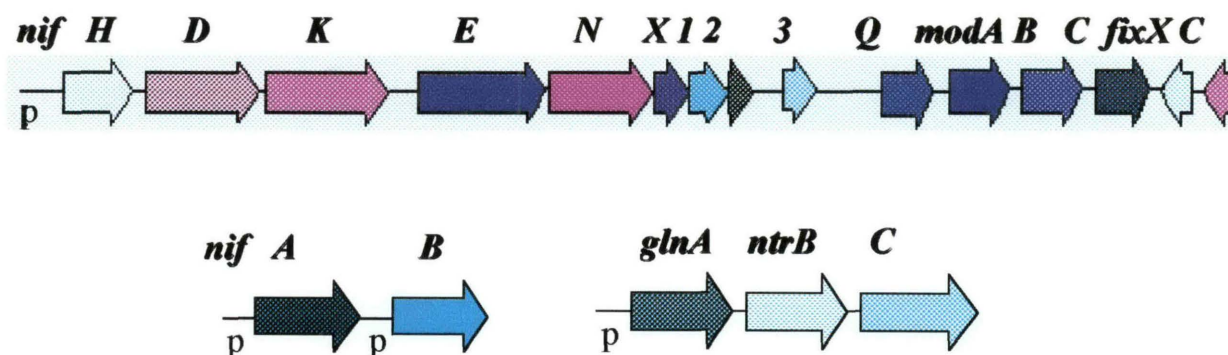


A importância atribuída a espécie *H. seropedicae* provém da sua capacidade de fixação de nitrogênio em associação com gramíneas de interesse econômico (DOBEREINER et al., 1992a).

O genoma de *H. seropedicae* possui aproximadamente 5,7 milhões de pares de bases (RAMOS, 2003) e regiões contendo os genes envolvidos com o metabolismo de nitrogênio já foram sequenciadas no Núcleo de Fixação de Nitrogênio (figura 5). A primeira região do genoma de *H. seropedicae* a ser sequenciada continha os genes *nifA* e *nifB* (SOUZA et al., 1990). O gene *glnB* que codifica para a proteína PII foi também isolado e sequenciado (BENELLI et al., 1997). Em uma terceira região sequenciada foram isolados os genes *glnAntrBC* e verificou-se que eles se encontram em um único operon (PERSUHN et al., 2000). A maior região sequenciada do genoma de *H. seropedicae* possui os genes *nifHD* e parte do gene *nifK* (MACHADO et al., 1995). Uma região de 5,1Kb a jusante dos genes *nifHDK* foi isolada e sequenciada (KLASSEN et al., 1999a). Os genes *nifENXorf1orf2orf3* foram identificados nesta região bem como os genes *nifQmodABCfixXC*. Além desses, foram isolados os genes *modE*, localizado na região a montante do gene *nifA* (VOIGT, 2000), e os genes *glnK* e *amtB* que codificam a proteína GlnK e o possível transportador de íons amônio, AmtB (NOINDORF, 2000).

O modelo atual sugere que a regulação da fixação de nitrogênio em *H. seropedicae* é modulada por amônio e oxigênio e envolve o sistema *ntr* e o sistema de transdução de sinal *nif* específico envolvendo as proteínas NifA e PII (PEDROSA et al., 2001). Mutantes *ntrC* e *ntrB* de *H. seropedicae* apresentaram baixa atividade de GS e deficiência nas atividades de adenililação/desadenililação de GS (PERSUHN et al., 2000), o que acaba por envolver o gene *glnE* e o seu produto, a proteína GlnE, na regulação da fixação de nitrogênio. Desse modo torna-se necessário a continuidade das pesquisas na caracterização dos genes envolvidos na fixação de nitrogênio.

**FIGURA 5: Organização estrutural dos genes envolvidos com a fixação biológica de nitrogênio em *H. seropedicae*.** (Autoria da figura: Giseli Klassen)



Genes envolvidos na fixação de nitrogênio em *H. seropedicae* sequenciados e caracterizados no Núcleo de Fixação Biológica de Nitrogênio do Departamento de Bioquímica da UFPR.

## 2. OBJETIVOS

### 2.1 Objetivos gerais

- Este projeto tem como objetivo geral a identificação, caracterização estrutural e mutagênese sítio dirigida do gene *glnE* de *Herbaspirillum seropedicae*.

### 2.2 Objetivos específicos

- Identificar, no banco de dados do programa GENOPAR, clones que contenham insertos de DNA de *H. seropedicae* correspondentes ao gene *glnE*;
- Sequenciar as extremidades dos insertos de DNA para confirmar a identidade dos clones;
- Analisar a sequência de nucleotídeos do gene *glnE* e sua vizinhança, através de programas de Bioinformática, identificando as regiões codificadoras de proteínas;
- Determinar a similaridade do gene isolado com genes *glnE* de outros organismos;
- Obter mutantes *glnE* plasmidiais por mutagênese sítio dirigida;
- Selecionar mutantes *glnE* cromossomais de *H. seropedicae*;
- Determinar a atividade da nitrogenase nos mutantes *glnE* de *H. seropedicae*.

### 3. MATERIAL E MÉTODOS

#### 3.1 Microrganismos e plasmídeos

As estirpes de bactérias e os plasmídeos utilizados estão listados no quadro abaixo.

Estirpes	Genótipo e/ou fenótipo	Referência
<i>E. coli</i>		
DH10B	$\phi 80\text{dlacZ}\Delta M15 \Delta\text{lacX}74 \text{ deoR } \text{recA1} \text{ endA1 } \text{araD139 } \Delta(\text{ara}, \text{leu})7697 \text{ galU } \text{galK } \lambda^- \text{ rpsL } \text{nupG}$	GRANT et al., 1990
<i>H. seropedicae</i>		
SMR1	$\text{Sm}^R, \text{nif}^+$	PEDROSA et al., 1997
HEL	$\text{Sm}^R, \text{Km}^R, \text{nif}^+, \text{glnE}^-$	Este trabalho
Plasmídeos		Referência
PUC18	$\text{LacZ}, \text{lacI}, \text{Amp}^R$	SAMBROOK et al., 1989
PUC4K	$\text{Amp}^R, \text{Km}^R$	Amersham Biosciences
PHF16	$\text{LacZ}, \text{lacI}, \text{Amp}^R, \text{Km}^R, \text{glnE}^-$	Este trabalho

#### 3.2 Condições e meios de cultivo das bactérias

As estirpes de *E. coli* foram cultivadas em meio líquido Luria-Bertani Broth (LB) ou Terrific Broth (TB) a 37°C e sob agitação (200 rpm). Para cultivo em meio sólido foi utilizado o meio Luria-Bertani Agar (LA) (SAMBROOK et al., 1989). Ao meio foram adicionados antibióticos, de acordo com a resistência da estirpe ou do plasmídeo que continham.

O meio Luria-Bertani Broth (LB) possui a seguinte composição:

Tryptona	10g/L
Extrato de levedura	5g/L
NaCl	10g/L

O meio LA sólido é preparado pela adição de 15 g/L de ágar ao meio líquido antes da autoclavagem.

O meio Terrific Broth (TB) possui a seguinte composição:

Bacto-triptona	12g/L
Extrato de levedura	24g/L
Glicerol	4mL/L

As estirpes de *H. seropedicae* foram cultivadas a 30°C, sob agitação a 130rpm em meio NFbHP (KLASSEN et al., 1997), utilizando malato como fonte de carbono, na presença dos respectivos antibióticos. O meio NFbHP possui a seguinte composição:

MgSO <sub>4</sub> .7H <sub>2</sub> O	2 x 10 <sup>-1</sup> g/L
NaCl	1 x 10 <sup>-1</sup> g/L
CaCl <sub>2</sub>	2 x 10 <sup>-2</sup> g/L
Ácido nitrilo-triacético	5,6 x 10 <sup>-2</sup> g/L
FeSO <sub>4</sub> . 7H <sub>2</sub> O	2 x 10 <sup>-2</sup> g/L
Malato de sódio	5 g/L
Biotina	1 x 10 <sup>-4</sup> g/L
Na <sub>2</sub> MoO <sub>4</sub> .2 H <sub>2</sub> O	2 x 10 <sup>-3</sup> g/L
MnSO <sub>4</sub> . H <sub>2</sub> O	2,35 x 10 <sup>-3</sup> g/L
H <sub>3</sub> BO <sub>3</sub>	2,8 x 10 <sup>-3</sup> g/L
CuSO <sub>4</sub> .5 H <sub>2</sub> O	8 x 10 <sup>-5</sup> g/L
ZnSO <sub>4</sub> .7 H <sub>2</sub> O	2,4 x 10 <sup>-4</sup> g/L



No momento do uso foram adicionados ao meio NFb 50mL/L de solução estéril de fosfatos (159,4g/L de  $\text{KH}_2\text{PO}_4$  e 17,8g/L de  $\text{K}_2\text{HPO}_4$ ) e 20mL/L de cloreto de amônio 1mol/L (NFbHPN) ou 5mL/L de glutamato de sódio 1mol/L (NFbHPG) também estéreis, conforme as condições necessárias para cada experimento. Os meios NFbHP sólido e semi-sólido foram preparados pela adição ao meio líquido de 15 e 1,75g/L de ágar, respectivamente.

### **3.2.1 Antibióticos**

Os antibióticos usados foram: ampicilina 250 $\mu\text{g/mL}$  e canamicina 50 $\mu\text{g/mL}$  para *E. coli*; para *H. seropedicae* foi utilizado 500 $\mu\text{g/mL}$  de canamicina e 80 $\mu\text{g/mL}$  de estreptomicina. As soluções estoque de estreptomicina, canamicina e ampicilina foram preparadas em água destilada e esterilizadas em membrana filtrante Millipore de 0,2 $\mu\text{m}$ . Os estoques foram mantidos a  $-20^\circ\text{C}$  e as alíquotas a  $4^\circ\text{C}$ .

## **3.3 Manipulação de DNA**

Técnicas de purificação de plasmídeos, ligação, digestão do DNA com endonucleases de restrição, eletroforese em gel de ágar e agarose, isolamento de fragmentos de DNA em gel de agarose de baixo ponto de fusão foram executadas segundo a literatura (SAMBROOK et al., 1989; AUSUBEL et al., 1984).

### **3.3.1 Purificação de DNA plasmidial**

A extração de plasmídeos a partir de estripes de *E. coli* foi realizada segundo SAMBROOK (1989). Alíquotas de 1,5mL da cultura cultivada até D.O.<sub>600</sub> ~2,0 foram centrifugadas a 13.000rpm por 1 minuto. As células foram, então, ressuspensas em 150 $\mu\text{L}$  de tampão GET (50 mmol/L glicose, 25mmol/L Tris-HCl

pH 8,8, 10 mmol/L EDTA pH8,0) e lisadas com 150µL de solução contendo NaOH 0,2 mol/L e SDS 1%, sob agitação manual. A esta mistura foram adicionados 150µL de KacF (acetato de potássio 3 mol/L, ácido fórmico 1,8 mol/L) e, após agitação, foi incubada em banho de gelo por 5 minutos. Foi adicionado, então, 100µL de uma mistura de fenol:clorofórmio:álcool isoamílico (25:24:1) e as amostras foram centrifugadas por 5 minutos a 13.000rpm. O sobrenadante resultante foi coletado e precipitado com 2 volumes de etanol absoluto. Depois de 20 minutos de repouso em temperatura ambiente as amostras foram centrifugadas a 13.000rpm por 10 minutos. O precipitado foi lavado com etanol 70% e seco à vácuo e, posteriormente, dissolvido em 20µL de água estéril.

### **3.3.2 Purificação de DNA genômico de *H. seropedicae***

Células de *H. seropedicae* de cerca de 1,5mL de cultura, crescida em meio NFbHP, foram coletadas por centrifugação a 13000rpm por 1 minuto. As células foram ressuspensas em 200 µL de tampão GET (50 mmol/L glicose, 25mmol/L Tris-HCl pH 8,8, 10 mmol/L EDTA pH8,0) e lisadas por tratamento com lisozima 100µg/mL a 50°C por 20 minutos. Em seguida, o material foi incubado com proteinase K 200ug/mL a 37°C por 12 horas. Após esse período a mistura foi centrifugada e o sobrenadante coletado foi homogeneizado lentamente com 1 volume de fenol:clorofórmio:álcool isoamílico e novamente centrifugado. O DNA genômico presente na fase aquosa foi precipitado pela adição de 2 volumes de etanol 96%, lavado com etanol 80%, seco e ressuspenso em 200µL de água ultra pura.

### **3.3.3 Seqüenciamento de DNA plasmidial**

O seqüenciamento de DNA foi realizado utilizando o método cujo princípio é descrito por SANGER et al. (1977). O sistema de reação foi feito utilizando-se aproximadamente 350ng de DNA de fita dupla, 10pmol de iniciador,

4µL de ET terminador mix (Amersham Biosciences) e água ultra pura suficiente para 20µL. O sistema foi amplificado em termociclador Eppendorf com as seguintes condições: 95°C 1 minuto e 35 ciclos de 94°C 20 segundos e 60°C 1 minuto e 30 segundos . A seguir o produto das reações foi transferido para um novo tubo contendo 7µL de acetato de amônio (7,5 mmol/L) e 70µL de etanol 96%. Após precipitação, a amostra de DNA foi centrifugada, lavada duas vezes com etanol 80% e seca à vácuo. O DNA foi ressuspenso, desnaturado a 95°C por 2 min e aplicado no gel do sequenciador automático.

### **3.4 Análise da seqüência do gene *glnE* de *Herbaspirillum seropedicae***

A análise da seqüência do gene *glnE*, bem como de sua vizinhança, foi realizada através de programas de Bioinformática, disponíveis em nosso laboratório ou de acesso público na Internet, e também manualmente.

As fases abertas de leitura (do inglês: *Open Reading Frame* – ORF) foram identificadas com o auxílio do programa FramePlot (ISHIKAWA e HOTTA). Seqüências de aminoácidos de proteínas homólogas à proteína GlnE de *H. seropedicae* foram obtidas no banco de dados do NCBI (<http://www.ncbi.nlm.nih.gov>). Essas seqüências foram alinhadas e comparadas para a identificação de regiões conservadas através do programa ClustalW (THOMPSON et al., 1994). A identificação de domínios estruturais foi feita com o auxílio do programa SMART (SCHULTZ et al., 2000).

A seqüência completa de nucleotídeos do gene *glnE* de *H. seropedicae* foi obtida a partir do banco de dados do programa GENOPAR. A identificação de possíveis regiões promotoras, sítio de ligação para ribossomo e possíveis seqüências terminadoras foram realizadas manualmente. A tradução da seqüência de nucleotídeos para a seqüência de aminoácidos foi feita pelo programa STRIDER (DOUGLAS, 1994).

### 3.5 Preparo de células eletrocompetentes

A estirpe DH10B de *E. coli* foi cultivada em meio LB até D.O.<sub>600</sub> entre 0,8 e 1,0. As células foram coletadas por centrifugação a 5.000rpm por 5 minutos a 4°C. O precipitado de células foi lavado 2 vezes com água ultra pura e gelada e uma vez com glicerol 15% estéril e gelado, sendo posteriormente ressuspensas em 500µL da mesma solução.

A estirpe SMR1 de *H. seropedicae* foi cultivada em meio NFbHP malato contendo NH<sub>4</sub>Cl 20mmol/L e 5g/L de solução fostato. O procedimento utilizado no preparo de células eletrocompetentes foi idêntico ao de *E. coli*.

### 3.6 Transformação bacteriana

Os plasmídeos foram transferidos para as diferentes estirpes por eletroporação. Aliquotas de 20µL de células eletrocompetentes (item 3.5) foram misturadas a alíquotas de 1µL do sistema de ligação (item 3.3) a ser eletroporado. A mistura foi transferida para uma cubeta de eletroporação (BRL) previamente resfriada e colocada na câmara de eletroporação. As amostras foram submetidas a um choque elétrico de 12 a 16kV para *E. coli* e de 8 a 10kV para *H. seropedicae*. As células de *E. coli* foram recuperadas em 1mL de meio Terrific Broth por um hora a 37°C. As células de *H. seropedicae* foram recuperadas por 3 horas em 1mL de meio NFbHP malato contendo NH<sub>4</sub>Cl 20mmol/L, 5g/L de solução fosfato e 1/10 do volume de meio LB.

### 3.7 Seleção de mutantes cromossomais *glnE* de *Herbaspirillum seropedicae*

A estirpe mutante de *H. seropedicae* foi obtida por inativação do gene *glnE* através da inserção de um cassete de canamicina (Km) obtido a partir do vetor pUC4K (Amersham Pharmacia Biotech).

### 3.7.1 Isolamento do cassete de canamicina

Para isolamento do cassete de canamicina, o vetor pUC4K (anexo1) foi digerido com a enzima de restrição *SaII* para liberação do fragmento correspondente ao cassete de canamicina (1,2kb). O fragmento de DNA de interesse foi extraído de um gel de agarose de baixo ponto de fusão 0,7%, conforme método descrito em SAMBROOK et al., (1989).

### 3.8 Mutagênese sítio dirigida

O plasmídeo HS03-FP-00-000-014-H11, identificado no banco de dados do Projeto GENOPAR, possui 1.700 pares de bases correspondentes à região intermediária do gene *glnE* e foi usado como alvo para inserção do cassete. Para isso foi realizado um ensaio onde alíquotas de um sistema de restrição foram coletadas nos períodos de 15, 30, 45 e 60 minutos com o objetivo de se obter uma digestão parcial do plasmídeo H11 com a enzima *SaII*. O plasmídeo digerido foi misturado, em um sistema de ligação, com o fragmento de DNA correspondente ao cassete de canamicina, para gerar clones mutantes para o gene *glnE*.

O plasmídeo mutante HS03-FP-00-000-014-H11 contendo inserto de canamicina foi eletroporado em *H. seropedicae* estirpe Smr1. Os mutantes *glnE* foram selecionados em NFbHPN sólido contendo 80µg/mL de estreptomicina e 500µg/mL de canamicina.

### 3.9 Hibridização do DNA

A transferência dos fragmentos de DNA para a membrana foi realizada pelo método de transferência por capilaridade ascendente como descrito em SAMBROOK et al., 1989. As amostras de DNA cromossômico foram tratadas com a enzima de restrição *EcoRI* seguido de separação em gel de agarose 0,8% em tampão TBE. Após a corrida, o gel foi corado com brometo de etídio, visualizado

em transiluminador de luz ultravioleta e registrado por fotografia. Após este procedimento, o gel foi tratado sucessivamente com solução de depuração (0,25mol/L de HCl) por 7 minutos, solução desnaturante (NaOH 0,5 mol/L; NaCl 1,5 mol/L) por 30 minutos, solução neutralizante (Tris-HCl 1 mol/L pH7,5/ NaCl 1,5 mol/L) por 20 minutos, e solução de transferência SSC 15X (SSC 20X - NaCl 3mol/L, citrato trissódico 0,3 mol/L pH 7,0) por 10 minutos.

As extremidades de uma tira de papel 3MM apoiada em uma lâmina de vidro foram imersas na solução de transferência SSC 15X. O gel foi então colocado com a face superior voltada para a ponte de papel 3MM e totalmente umedecida com solução SSC 20X. Uma membrana de nylon (Hybond, Armersham) do mesmo tamanho do gel e marcada na face de transferência foi colocada sobre a face inferior de gel. As bolhas de ar eventualmente presente entre a membrana e o gel foram cuidadosamente removidas. A membrana foi então coberta com 3 folhas de papel 3MM. O sistema foi montado em cuba contendo uma camada de tampão SSC 15X de aproximadamente 1cm e selado com filme de plástico, permitindo o fluxo de tampão somente através do gel apoiado na ponte de papel 3MM. Em seguida, uma camada de papel absorvente de aproximadamente 10cm foi colocada sobre a membrana, pressionada por lâminas de vidro pesando 300-500g. Os fragmentos de DNA foram transferidos do gel para a membrana de nylon por ação capilar ascendente imposta pelo papel absorvente, por 48 horas.

Após a transferência, o sistema foi desmontado e a membrana seca a temperatura ambiente. O DNA foi fixado à membrana por irradiação com luz ultravioleta (302nm) por 5 min.

### **3.10 Determinação da atividade da nitrogenase**

A atividade de nitrogenase foi determinada pelo método de redução do acetileno (DILWORTH, 1966; SCHÖLLORN & BURRIS, 1967) em estirpes de *H.*

*seropedicae* cultivadas em meio NFbHP semi-sólido. A atividade nitrogenase foi expressa em  $\eta$ mol de etileno formado por mg de proteína por minuto.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Identificação de clones no banco de dados do Programa GENOPAR contendo o gene *glnE* de *Herbaspirillum seropedicae*

O Programa GENOPAR tem o objetivo de seqüenciar completamente o genoma da bactéria fixadora de nitrogênio *Herbaspirillum seropedicae*. Para tanto foram construídas várias bibliotecas genômicas em *E. coli*, utilizando-se os vetores de clonagem pUC18 e pUC19 (SAMBROK et al., 1989). O banco genômico contém insertos de DNA de *H. seropedicae* com tamanhos entre 1,5 e 3,0 kb e as seqüências de DNA correspondentes às extremidades 5' e 3' destes insertos estão armazenadas no Banco de Dados do Programa GENOPAR ([www.genopar.org.br](http://www.genopar.org.br)). Fazem parte deste Banco de Dados também os resultados das análises feitas pelo programa BLASTX das seqüências do Programa GENOPAR contra o banco de proteínas do NCBI.

Uma vez que este projeto teve por objetivo caracterizar o gene *glnE* de *H. seropedicae*, foi realizada uma busca no banco de dados do Programa GENOPAR utilizando a palavra-chave *glnE* (figura 6). Como resultado foram identificados três clones: HS03-FP-00-000-014-H11, HS08-FP-00-000-005-E09 e HS12-FP00-000-001-D09. Entre estes, o clone HS03-FP-00-000-014-H11 foi selecionado para ensaios de mutagênese sítio dirigida e, a partir deste ponto, este clone será referido apenas com H11. O clone H11 foi selecionado pois continha uma região intermediária significativa do gene *glnE* com aproximadamente 1700 pares de bases e também um sítio único para a enzima *SalI* que possibilitou os experimentos subseqüentes.

A seqüência obtida com o iniciador universal e disponível no banco de dados do Programa GENOPAR possui 900 pares de bases e apresenta grande similaridade com a porção inicial do gene *glnE* (figura 7a). A comparação da seqüência obtida com o iniciador reverso também indica similaridade (figura 7b)



com a porção inicial do gene *glnE*. Isso se explica pelo fato das extremidades N-terminal e C-terminal da proteína GlnE serem semelhantes em sua sequência de aminoácidos (van HEESWIJK *et al.*, 1993). O DNA inserto do plasmídeo H11 foi novamente seqüenciado com os iniciadores universal e reverso para confirmação da identidade (figura 8).

Segundo dados de anotação do genoma do Programa GENOPAR o gene *glnE* de *H. seropedicae* possui 2.784 pares de bases. O inserto de DNA de *H. seropedicae* no plasmídeo H11 possui aproximadamente 1.700 pares de bases e a sequência obtida com o iniciador universal (figura 7a) apresenta similaridade com a região do aminoácido 109 até o aminoácido 330 da proteína GlnE. Logo, pode-se inferir que o plasmídeo corresponde à região intermediária do gene *glnE* de *H. seropedicae*.

FIGURA 6: Página de pesquisa do Programa GENOPAR

**GENOPAR**  
**Pesquisa BLAST**

**Read**

**Descrição**

**Organismo**

**Banco**

**Registros**

Procurar Configs na montagem: ☐ Ultima ☐ Anotacao ☒ \*Nao procurar\*

[\(Instruções de Pesquisa\)](#)

Exibindo 31 registros encontrados

ID	IDENTIFICAÇÃO DA SEQUÊNCIA	DESCRIÇÃO	ORGANISMO	VALOR E
848804	<a href="#">HS09-RN-00-000-065-E08.g</a>	(NC_003197) regulatory protein, P-II 2, for nitrogen assimilation by glutamine synthetase, regulates GlnL (NRII) and GlnE (ATase)	Salmonella typhimurium LT2	9e-54
14049	<a href="#">HS03-FP-00-000-014-H11.b</a>	GLNE_HAEIN GLUTAMATE-AMMONIA-LIGASE ADENYLYLTRANSFERASE (GLUTAMINE-SYNTHETASE ADENYLYLTRANSFERASE) (ATASE)		1e-44
14051	<a href="#">HS03-FP-00-000-014-H11.b</a>	glutamate-ammonia-ligase adenylyltransferase (glnE)	Haemophilus influenzae Rd	1e-44
	<a href="#">HS03-FP-00-000-014-H11.b</a>	GLNE ECOLI GLUTAMATE-AMMONIA-LIGASE		

A palavra chave *glnE* foi colocada no campo descrição da página de pesquisa (retângulo verde). A busca no banco de dados do Programa GENOPAR indicou que uma sequência similar ao gene *glnE* havia sido identificada no plasmídeo HS03-FP-00-000-014-H11 (retângulo vermelho).

**FIGURA 7: Comparação da sequência do plasmídeo HS03-FP-00-000-014-H11, obtidas no pelo Programa GENOPAR, com o GenBank.**

**a) Sequência obtida com o iniciador universal**

Probable Glutamine-Synthetase Adenylyltransferase - *Ralstonia solanacearum*

Length = 955 E-value= 2e-64

```
Query: 50 LDEVVTTMTGFADFAVQTHLAALMQEQTALYGLPIGEESGRPQEMIVLGMGKLGGGELNV 229
      LDEV MT FA+FAV+ +A + QE AL+G P+G+ SG QE++V+GMGKLGG ELNV
Sbjct: 109 LDEVTGAMTDFAEFVRAAIAVIGQELGALHGQPVGQSSGEVQELVVVGMGKLGGRELV 168

Query: 230 SSDIDLIFVYPEDGDTRAEAGQKSLSNHEFFVRLGKKLIGALAEITEDGFTFRVDMALRP 409
      SSDIDLIF+Y E+GDT + G + LSNHE+F +LG++LI ALA++TEDG+ FRVDM LRP
Sbjct: 169 SSDIDLIFLYDEEGDT--QGGPRPLSNHEYFTKLGRRLINALADVTEDGYVFRVDMRLRP 226

Query: 410 NGNSGPLVASFNMVEEYLVRQGREWERYAWTKARALT--GTPED---IATLEAISRPFFIF 574
      NG++GPL S M+EEY V QGREWERYAW K R +T G+P I LE ++ PF+F
Sbjct: 227 NGDAGPLACSLGMLEEFVVGQGREWERYAWIKGRVITDPGSPHAARVIQQLERVTTTPFVF 286

Query: 575 RRYLDYGSIDALRSMHGQIRAEVKRHEALLSDRRKTSSWGRADSGEQVH 721
      RRYLDYG I A+R++H QIRAE + + + G DS Q H
Sbjct: 287 RRYLDYGVIAAIRALHAQIRAEAAKRSNAANRQH-----GGHDSQVQAH 330
```

**b) Sequência obtida com o iniciador reverso**

Glutamine-synthetase adenylyltransferase - *Haemophilus influenzae*

Length = 981 E-value = 4e-04

```
Query: 150 GLPIGEESGRPQEMIALGMGKLGVGELDVSSDIERGGVYPEVRRRIACAEVRGTNSGAPHQ 329
      G P+ +E G PQ++ LGMGKLG EL+ SSDI+ YP V TN+
Sbjct: 160 GTPM-DEQGNPQQLYILGMGKLGGFELNFSSDIDLIFTYPSQGE---TSVENTNARRSVD 215

Query: 330 *SSFAAELGQ 359
      F LGQ
Sbjct: 216 NGKFFTRLGQ 225
```

Sequências do plasmídeo HS03-FP-00-000-014-H11 obtidas com os iniciadores universal (a) e reverso (b). E-Value é um índice que reflete a probabilidade da sequência submetida ser igual à encontrada no banco aleatoriamente. Quanto menor o “E Value”, maior a probabilidade de as sequências serem semelhantes em decorrência de origem evolutiva comum ou mesma função (o baixo valor de E-Value obtido indica que a probabilidade dessa similaridade ter sido encontrada por acaso é muito baixa). Na comparação entre as sequências “Query” representa o intervalo de bases da sequência submetida que apresenta similaridade com a sequência depositada no GenBank. “Subject” representa a sequência de aminoácidos correspondente ao intervalo de bases analisado. “Length” refere-se ao número de aminoácidos da proteína com a qual foi encontrada similaridade. O programa utilizado foi o BLASTx (ALTSCHUL et al., 1997).

## FIGURA 8: Confirmação da identidade do plasmídeo H11

Glutamine-synthetase adenylyltransferase - *Haemophilus influenzae*

Length = 981

E-value 2e-57

```

Query: 48  QRLEHDLAS---GLSLPAAMRRRLRNLVICTLITRDLDGRGDLDEVVTTMTGFADFAVQTH 218
          +RL  +LA+      +L  +R+ RN+ +  L              ++E+  ++  A+  +
Sbjct: 88  ERLNKELANVHDETALYQTLRQFRNIEMAKLSFCQSLNLATVEEIFVQLSQAESLIIAA 147

Query: 219 LAALMQEQTALYGLPIGEESGRPQEMIVLGMGKLGGGELNVSSDIDLIFVYPEDGDTRAE 398
          L  +  A  G P+ +E G PQ++ +LGMGKLGG ELN SSDIDLIF YP  G+T  E
Sbjct: 148 RDWLYHQACAEIGTPM-DEQGNPQQLYILGMGKLGGFELNFSSDIDLIFTYPSQGETSVE 206

Query: 399 A--GQKSLSNHEFFVRLGKKLIGALAEITEDGFTFRVDMALRPNGNSGPLVASFNMVEEY 572
          ++S+ N +FF RLG++LI AL E T DGF +R DM LRP G+SG LV SF+ +E+Y
Sbjct: 207 NTNARRSVDNGKFFTRLGQRLISALDEFTSDGFFVYRTDMRLRPFGDSGVLVLSFSAMEQY 266

Query: 573 LVRQGREWERYAWTKARALTGTPED--IATLEAISRPFFIFRRYLDYGSIDALRSMHGQIR 746
          QGR+WERYA  K R L    +D  ++ L+ + RPF++RRY+D+  I +LR M G+I
Sbjct: 267 YQDQGRDWERYAMIKGRVLGAQTKDPNVSKLKKMLRPFVYRRYIDFSVIQSLREMKGKIE 326

```

Seqüência do clone H11 obtida com o iniciador universal.

#### 4.2 Análise de um fragmento de DNA cromossomal de *H. seropedicae* contendo o gene *glnE*

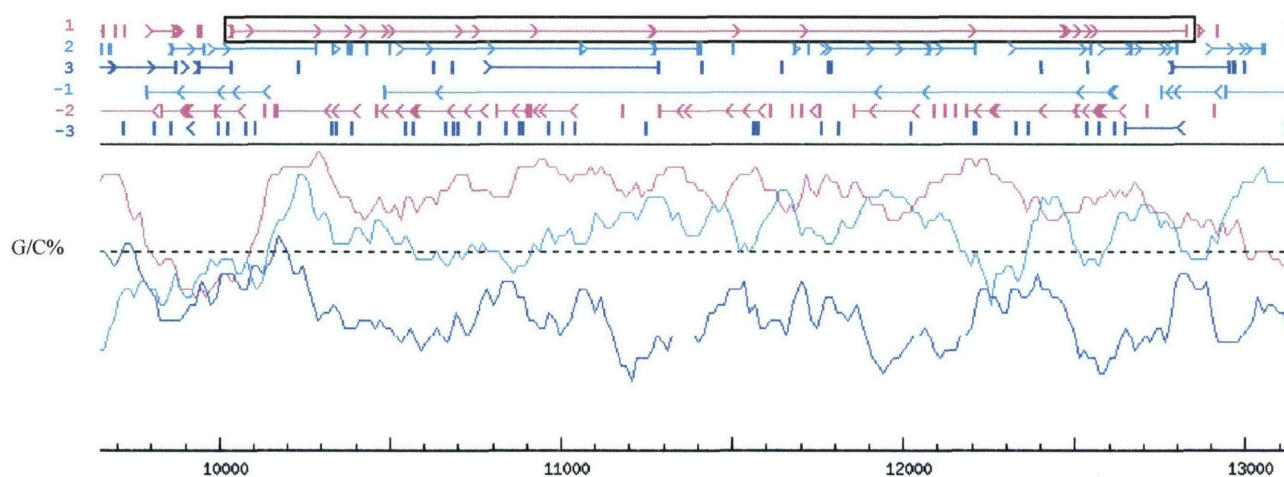
A região do genoma de *H. seropedicae* na qual esta inserido o gene *glnE* foi completamente seqüenciada (Programa GENOPAR). Um fragmento de 25kb, contendo o gene *glnE*, foi selecionado dessa região para uma análise mais detalhada.

Utilizando-se o programa FramePlot (ISHIKAWA e HOTTA, 1999) foram identificadas as prováveis fases de leitura (ORFs) contidas na sua seqüência. Esse programa gera, a partir da seqüência submetida, dois gráficos: um contendo as possíveis fases apresentadas em formato de setas, dispostas nas seis fases de leitura - 3, -2, -1, 1, 2 e 3, e outro indicando a porcentagem de G-C na terceira base do códon e para cada fase de leitura separadamente de cada ORF (figura 9). Cada seta permite acessar a seqüência de bases da ORFs, bem como a seqüência de aminoácidos. A validação das ORFs se deu pela comparação da seqüência com o banco de dados de proteínas do NCBI (<http://www.ncbi.nlm.nih.gov>) através do programa BLASTp (ALTSCHUL et al., 1997) que faz uma busca por de similaridade entre seqüências de aminoácidos.

Um total de 17 genes foram identificados no fragmento analisado (figura 10). A montante ao gene *glnE* foi identificado um gene que codifica uma provável proteína transmembrana de função desconhecida e a jusante foi identificado um gene que codifica uma metil-transferase dependente de SAM (S-adenosil metionina). Em *E. coli* esta proteína esta relaciona com a resistência a telurite ( $\text{TeO}_3^{-2}$ ) (LIU, et al, 2000). Ambos são expressos em orientação contrária ao gene *glnE* o que sugere que o mesmo é monocistrônico. Os demais genes identificados pelo programa BLAST apresentaram alta similaridade para as seguintes proteínas: 1) UDP-N-acetilenolpiruvoil glucosamina redutase; 2) Proteína conservada não caracterizada; 3) Fosfo-2-hidro-3 deoxiheptonato aldolase; 4) Provável proteína TDLL; 5) Amido hidrolase; 6) Proteína hipotética CV2975; 7)

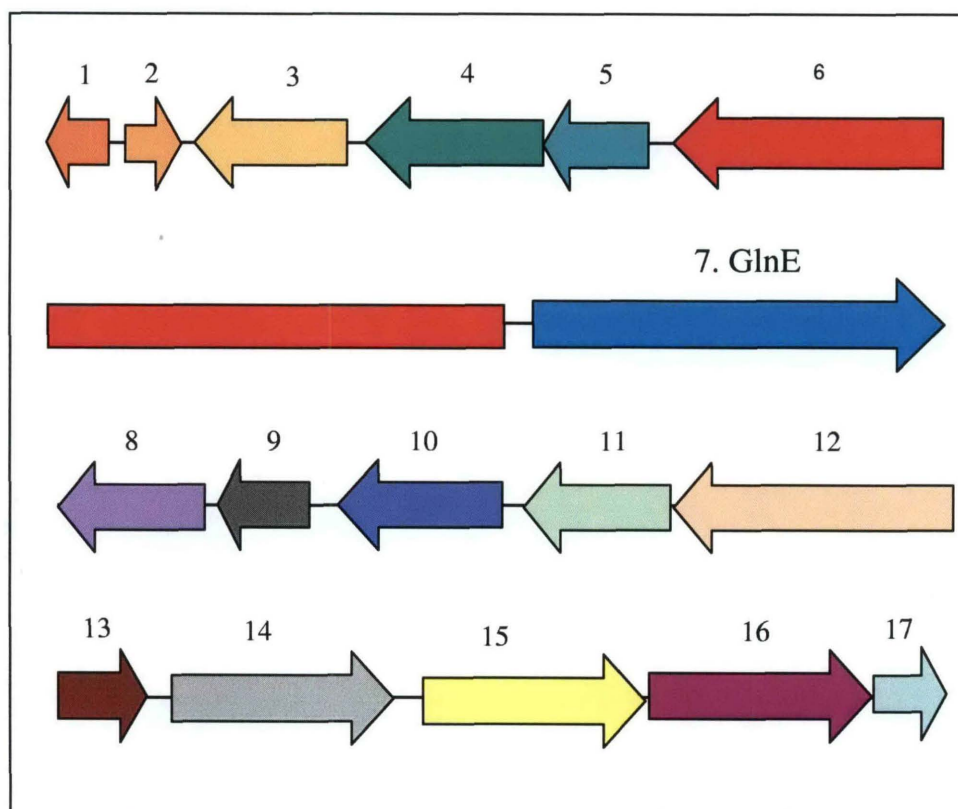
Gliceraldeído 3 fosfato desidrogenase; 18) Acetil transferase; 19) Transcetolase; 10) Proteína conservada, 11) Amidotransferase; 12) Transportador ABC lipoproteína; 13) Transportador ABC permease; 14) Proteína hipotética.

**Figura 9: Identificação de ORFs pelo programa FramePlot**



As setas representam as ORFs que foram localizadas pelo programa FramePlot. Os números à esquerda indicam as fases de leitura 1, 2, 3, -1, -2 e -3. As linhas do gráfico abaixo das setas referem-se ao conteúdo de G/C de cada ORF. O gene *glnE* está destacado no box.

**FIGURA 10: Identificação de proteínas a montante e a jusante ao gene *glnE* em *H. seropedicae*.**



- 1) UDP-N-acetilenolpiruvato glucosamina redutase; 2) Proteína conservada não caracterizada; 3) Fosfo-2-hidro-3 deoxiheptonato aldolase; 4) Provável proteína TDLL; 5) Amido hidrolase; 6) Provável proteína de membrana; 7) *GlnE*; 8) Metil transferase SAM dependente; 9) Proteína hipotética CV2975; 10) Gliceraldeído 3 fosfato desidrogenase; 11) Acetil transferase; 12) Transcetolase; 13) Proteína conservada; 14) Amidotransferase; 15) Transportador ABC lipoproteína; 16) Transportador ABC permease; 17) Proteína hipotética.



#### 4.2.1 Análise do gene *glnE* de *H. seropedicae*

O gene *glnE* de *H. seropedicae* possui 2.784 pares de bases que codificam a proteína GlnE de 927 aminoácidos (figura 11). A montante do códon de início de transcrição (5'ATG3') há um provável sítio de ligação para ribossomo (5'GAAG3') e três possíveis regiões promotoras: uma dependente do fator  $\sigma^{70}$  e duas dependentes do fator  $\sigma^{54}$ . A comparação entre as seqüências consenso para esses promotores e as seqüências localizadas a montante ao códon de início de transcrição do gene *glnE* apresentam uma baixa similaridade (figura 12). Além disso, um desses possíveis promotores pode estar relacionado com a proteína transmembrana localizada a montante ao *glnE* e que é transcrita no sentido inverso em relação a esse gene (figura 10). Existe uma probabilidade maior de que o promotor realmente envolvido na transcrição do gene *glnE* seja aquele dependente de  $\sigma^{70}$ , já que este gene é constitutivamente expresso e não apenas em condições limitantes de nitrogênio. Para identificar sob influência de qual desses promotores o gene *glnE* está sujeito, seria necessário construir uma fusão da região promotora com um gene repórter e medir a sua atividade. A jusante ao códon de término de transcrição (5'TGA3') há uma seqüência palindrômica (5'GAAAAACGGGATGCACctgccGGTGTATCCCGTTTTTC3') que pode estar atuando como um terminador de transcrição pela formação de um grampo (independente de  $\rho$ ). Em *E. coli* o gene *glnE* é co-transcrito em um operon com os genes ... (van Heeswijk WC)

#### 4.2.2 Análise da proteína GlnE

A proteína GlnE de *H. seropedicae* foi comparada com proteínas GlnE de outros organismos. A identificação de regiões conservadas foi determinada pelo alinhamento múltiplo de seqüências através programa ClustalW (THOMPSON *et*



al., 1994). A análise dos domínios da proteína foi feita utilizando-se o programa SMART (SCHULTZ et al., 2000).

GlnE é uma enzima bifuncional que catalisa tanto a adenililação quanto a desadenililação da glutamina sintetase regulando sua atividade. Segundo JAGGI e colaboradores (1997) em *E. coli* esses domínios são homólogos estando o sítio de desadenililação na região N-terminal (aminoácidos 117 a 184) e o sítio de adenililação na região C-terminal (aminoácidos 643 a 711). O alinhamento múltiplo das seqüências de aminoácidos das proteínas GlnE realizado pelo programa ClustalW corrobora esse dado (figura 13). Na região N-terminal há um domínio conservado que é homólogo a outro domínio conservado presente na região C-terminal. Em *H. seropedicae* esses domínios podem ser representados pelas seqüências de aminoácidos:

- N-terminal: 151 M I V L G M G K L G G G E L N V S S D I D L I F V Y 177
- C-terminal: 674 F A V I A Y G K L G G K E L G Y A S D L D V V F L Y 700

Provavelmente os sítios de adenililação e desadenililação possuem mais aminoácidos do que os mostrados acima, mas essa é uma região altamente conservada e que possui maior homologia entre os domínios, apresentando aminoácidos iguais e outros que pertencem a um mesmo grupo. A jusante ao sítio de desadenililação, ainda na porção N-terminal, existem mais um região altamente conservada, mas esta não possui correspondente na porção C-terminal. Essa região conservada, intermediária entre os sítios catalíticos, pode estar relacionada ao sítio de interação da proteína GlnE com PII/PII-UMP (figura 13).

A busca por domínios estruturais permitiu a observação de dois domínios repetidos, um próximo ao início e outro próximo ao fim da proteína que, provavelmente, correspondem aos sítios de adenililação e desadenililação respectivamente (figura 14a). A busca por domínios conservados nas proteínas GlnE de *Ralstonia solanacearum* (figura 14b) e *Escherichia coli* (figura 14c) forneceram

resultados semelhantes, nos quais apenas dois domínios conservados foram reconhecidos. Este resultado corrobora os dados obtidos com o programa ClustalW. Observa-se que em *R. solanacearum* e *E. coli* os domínios repetidos são identificados como pertencentes à proteína GlnE, pois o genoma dessas bactérias já foi plenamente seqüenciado e anotado sendo que muitas das suas proteínas já foram caracterizadas. Entretanto, na seqüência de *H. seropedicae*, foram encontrados apenas dois domínios repetidos, característico de enzimas nucleotidil-transferase, mas que não foram identificados como GlnE. Isso pode estar relacionado com o fato da proteína ainda não ter sido caracterizada em *H. seropedicae*. O provável domínio correspondente ao sítio de ligação para PII/PII-UMP não aparece no resultado, possivelmente por ainda não ter sido descrito e não estar nos bancos de dados internacionais.

**FIGURA 11: Sequência do gene *glnE* de *H. seropedicae***

1 CAG GAT GCA ATG GCA GCT CGG GAT TCA AAG ATT TCT GGC AAT ACG GCA GTT CAG GCA GGT 60

61 GAA <sup>-24</sup> GCG GTG ACA TGA TCG CAT <sup>-12</sup> GGA AAA CAT GGC CAT GCC ATC TGC GCA ATG CCC TTC CTG 121

121 <sup>-24</sup> GGT GCA AAA TGG CCG <sup>-12</sup> <sup>-35</sup> ATA AAT TGA AAA ATG AAG GAT AGA ATC <sup>-10</sup> CCC TCT CTG GCC CAT CCG 181

181 TCA GGC CCA TGT CAG CCC TGT CCC ACC CTG TCC GAT TAT CGA CGC GGC GCA ACA TGA **CTG** 240

241 **AAG** CGA <sup>M</sup> CAT <sup>A</sup> GGC <sup>A</sup> CGC <sup>G</sup> CGG <sup>F</sup> TTT <sup>P</sup> TCC <sup>S</sup> TTC <sup>F</sup> TTT <sup>T</sup> TAC <sup>L</sup> ACT <sup>C</sup> TTG <sup>C</sup> CTG <sup>G</sup> CGG <sup>A</sup> AGC <sup>F</sup> TTT <sup>Q</sup> CCA <sup>T</sup> GAC <sup>L</sup> GCT 300

301 <sup>M</sup> TAT <sup>P</sup> GCC <sup>V</sup> GGT <sup>F</sup> CTT <sup>S</sup> CTC <sup>P</sup> GCC <sup>A</sup> TGC <sup>H</sup> CCA <sup>H</sup> TCA <sup>R</sup> CCG <sup>S</sup> CTC <sup>L</sup> CCT <sup>V</sup> CGT <sup>E</sup> CGA <sup>S</sup> GGC <sup>S</sup> CTC <sup>S</sup> GTC <sup>R</sup> CTC <sup>Y</sup> GCG CTA 360

361 <sup>Y</sup> CTA <sup>T</sup> CAC <sup>R</sup> CCG <sup>W</sup> CTG <sup>T</sup> GAC <sup>Q</sup> CCA <sup>A</sup> GGC <sup>D</sup> CGA <sup>P</sup> CCC <sup>A</sup> GGC <sup>R</sup> CCG <sup>P</sup> CCC <sup>A</sup> CGC <sup>L</sup> CCT <sup>V</sup> GGT <sup>A</sup> GGC <sup>A</sup> TGC <sup>A</sup> GCG <sup>A</sup> GGC GGC 420

421 <sup>A</sup> CGC <sup>G</sup> CGG <sup>L</sup> TGC <sup>D</sup> GCT <sup>R</sup> TGA <sup>F</sup> CCG <sup>L</sup> CGC <sup>Q</sup> CTT <sup>Q</sup> CCT <sup>R</sup> GCA <sup>R</sup> GCA <sup>L</sup> GCG <sup>E</sup> CCT <sup>H</sup> GGA <sup>D</sup> GCA <sup>L</sup> CGA <sup>S</sup> CCT <sup>G</sup> GGC <sup>A</sup> CAG <sup>S</sup> CGG 480

481 <sup>L</sup> CCT <sup>S</sup> GAG <sup>P</sup> CCT <sup>A</sup> GCC <sup>A</sup> GGC <sup>M</sup> GGC <sup>R</sup> CAT <sup>R</sup> GCG <sup>L</sup> CCG <sup>R</sup> GCT <sup>N</sup> GCG <sup>L</sup> CAA <sup>V</sup> CCT <sup>I</sup> GGT <sup>C</sup> CAT <sup>T</sup> CTG <sup>L</sup> CAC <sup>I</sup> CCT <sup>T</sup> GAT CAC 540

541 <sup>R</sup> GCG <sup>D</sup> CGA <sup>D</sup> CCT <sup>G</sup> CGA <sup>G</sup> TGG <sup>D</sup> CCG <sup>L</sup> TGG <sup>D</sup> CGA <sup>E</sup> CCT <sup>V</sup> CGA <sup>V</sup> AGT <sup>T</sup> CGT <sup>T</sup> AAC <sup>T</sup> TAC <sup>M</sup> CAT <sup>T</sup> GAC <sup>G</sup> CGG <sup>F</sup> CTT <sup>A</sup> TGC 600

601 <sup>D</sup> CGA <sup>F</sup> CTT <sup>V</sup> TGC <sup>Q</sup> CGT <sup>T</sup> GCA <sup>H</sup> GAC <sup>L</sup> CCA <sup>A</sup> CCT <sup>L</sup> GGC <sup>A</sup> GGC <sup>M</sup> GCT <sup>Q</sup> GAT <sup>E</sup> GCA <sup>Q</sup> GGA <sup>T</sup> ACA <sup>A</sup> GAC <sup>L</sup> TGC <sup>Y</sup> GCT <sup>G</sup> CTA CGG 660

661 <sup>L</sup> CCT <sup>P</sup> GCC <sup>G</sup> CAT <sup>E</sup> CGG <sup>E</sup> CGA <sup>S</sup> GGA <sup>G</sup> ATC <sup>R</sup> AGG <sup>P</sup> ACG <sup>Q</sup> GCC <sup>E</sup> GCA <sup>M</sup> GGA <sup>I</sup> AAT <sup>V</sup> GAT <sup>L</sup> CGT <sup>G</sup> GCT <sup>M</sup> GGG <sup>K</sup> CAT GGG CAA 720

721 <sup>L</sup> GCT <sup>G</sup> GGG <sup>G</sup> TGG <sup>E</sup> GGG <sup>N</sup> CGA <sup>V</sup> ACT <sup>S</sup> CAA <sup>S</sup> TGT <sup>D</sup> CTC <sup>I</sup> CTC <sup>D</sup> CGA <sup>I</sup> CAT <sup>D</sup> CGA <sup>F</sup> CCT <sup>V</sup> CAT <sup>Y</sup> CTT <sup>P</sup> CGT <sup>E</sup> CTA TCC CGA 780

781 <sup>D</sup> AGA <sup>G</sup> CGG <sup>T</sup> CGA <sup>R</sup> TAC <sup>A</sup> CCG <sup>E</sup> CGC <sup>A</sup> CGA <sup>G</sup> GGC <sup>Q</sup> CGG <sup>K</sup> CCA <sup>S</sup> GAA <sup>L</sup> ATC <sup>S</sup> GCT <sup>N</sup> GTC <sup>H</sup> CAA <sup>E</sup> CCA <sup>F</sup> TGA <sup>F</sup> ATT <sup>V</sup> CTT CGT 840

841 <sup>R</sup> GCG <sup>L</sup> GCT <sup>K</sup> GGG <sup>K</sup> CAA <sup>L</sup> GAA <sup>I</sup> GCT <sup>G</sup> CAT <sup>A</sup> CGG <sup>L</sup> TGC <sup>A</sup> GCT <sup>E</sup> GGC <sup>I</sup> CGA <sup>T</sup> GAT <sup>T</sup> CAC <sup>D</sup> TGA <sup>G</sup> GGA <sup>F</sup> CGG <sup>T</sup> TTT <sup>F</sup> CAC CTT 900

901 <sup>R</sup> CCG <sup>V</sup> GGT <sup>M</sup> CGA <sup>A</sup> CAT <sup>L</sup> GGC <sup>R</sup> GCT <sup>P</sup> GCG <sup>N</sup> CCC <sup>G</sup> CAA <sup>N</sup> CGG <sup>S</sup> CAA <sup>G</sup> CTC <sup>G</sup> CCG <sup>P</sup> CCC <sup>L</sup> GCT <sup>V</sup> GGT <sup>A</sup> GGC <sup>S</sup> CAG <sup>N</sup> TTT CAA 960

961 <sup>M</sup> CAT <sup>V</sup> GGT <sup>E</sup> GGA <sup>Y</sup> GGA <sup>L</sup> ATA <sup>R</sup> CCT <sup>Q</sup> GGT <sup>G</sup> GCG <sup>R</sup> CCA <sup>E</sup> GGG <sup>W</sup> CCG <sup>E</sup> AGA <sup>R</sup> ATG <sup>E</sup> GGA <sup>Y</sup> ACG <sup>A</sup> CTA <sup>T</sup> TGC <sup>K</sup> CTG GAC CAA 1020

1021 <sup>A</sup> GGC <sup>R</sup> GCG <sup>L</sup> CGC <sup>T</sup> CCT <sup>T</sup> CAC <sup>P</sup> CGG <sup>D</sup> CAC <sup>I</sup> GCC <sup>A</sup> GGA <sup>T</sup> AGA <sup>A</sup> CAT <sup>L</sup> CGC <sup>E</sup> CAC <sup>I</sup> GCT <sup>S</sup> GGA <sup>R</sup> AGC <sup>P</sup> GAT <sup>R</sup> CAG <sup>P</sup> CCG CCC 1080

1081 <sup>F</sup> CTT <sup>I</sup> CAT <sup>R</sup> CTT <sup>R</sup> CCG <sup>Y</sup> CCG <sup>L</sup> CTA <sup>D</sup> TCT <sup>G</sup> CGA <sup>S</sup> CTA <sup>I</sup> CGG <sup>D</sup> CTC <sup>A</sup> CAT <sup>L</sup> CGA <sup>R</sup> CGC <sup>S</sup> GCT <sup>M</sup> GCG <sup>H</sup> CTC <sup>G</sup> CAT GCA CGG 1140

1141 <sup>Q</sup> GCA <sup>I</sup> GAT <sup>R</sup> CCG <sup>A</sup> CGC <sup>E</sup> CGA <sup>V</sup> GGT <sup>K</sup> CAA <sup>R</sup> GCG <sup>Q</sup> CCA <sup>E</sup> GGA <sup>A</sup> AGC <sup>L</sup> CCT <sup>H</sup> GCA <sup>P</sup> CCC <sup>D</sup> GGA <sup>S</sup> CCG <sup>N</sup> CAG <sup>N</sup> CAA <sup>V</sup> CGT 1200

1201 <sup>K</sup> CAA <sup>L</sup> GCT <sup>G</sup> GGG <sup>G</sup> CCG <sup>I</sup> TGG <sup>R</sup> CGG <sup>E</sup> CAT <sup>I</sup> TCG <sup>E</sup> CGA <sup>F</sup> GAT <sup>T</sup> CGA <sup>T</sup> GTT <sup>S</sup> CAC <sup>V</sup> CAG <sup>Q</sup> CCA <sup>F</sup> GGT <sup>L</sup> ATT <sup>I</sup> CCA GCT GAT 1260

1261 <sup>R</sup> CCG <sup>G</sup> GGG <sup>R</sup> CGG <sup>D</sup> GCG <sup>A</sup> CGA <sup>L</sup> TGC <sup>R</sup> CGA <sup>D</sup> GCT <sup>R</sup> GCG <sup>S</sup> CGA <sup>T</sup> CCG <sup>T</sup> CTC <sup>T</sup> CAC <sup>T</sup> CCG <sup>T</sup> CAC <sup>L</sup> CCT <sup>R</sup> GCG <sup>T</sup> CAC CCT 1320

1321 <sup>A</sup> GGC <sup>A</sup> CGC <sup>K</sup> CAA <sup>D</sup> GGA <sup>L</sup> CCT <sup>L</sup> GCT <sup>A</sup> CGC <sup>P</sup> ACC <sup>V</sup> CGA <sup>V</sup> GGT <sup>E</sup> GGA <sup>Q</sup> GCA <sup>L</sup> ATT <sup>L</sup> GCT <sup>H</sup> GCA <sup>Y</sup> CGC <sup>T</sup> CTA <sup>F</sup> CAC CTT 1380

Continua



L R N L E H R L Q Y L E D A Q T H T L P  
 1381 CCT GCG CAA TCT CGA ACA CCG CCT GCA ATA CCT GGA AGA CGC CCA GAC CCA CAC CCT GCC 1440  
 V N P D D L L L V A R M M G Y E E P A A  
 1441 CGT CAA TCC CGA CGA CCT GCT GCT GGT GGC CCG GAT GAT GGG CTA TGA GGA ACC GGC CGC 1500  
 L L H E L E R Q R A I V A A Q F D A I F  
 1501 GCT GCT GCA CGA ACT GGA GCG CCA GCG CGC CAT CGT CGC CGC CCA GTT CGA CGC CAT CTT 1560  
 A D K Q S E A D N D G P S V S V S D N D  
 1561 CGC CGA CAA GCA GTC CGA GGC CGA CAA CGA CGG CCC CAG CGT GAG TGT GAG CGA TAA CGA 1620  
 N L E G L A D A L R L V G F P E A D I A  
 1621 CAA TCT CGA AGG ACT CGC CGA CGC CCT GCG CCT GGT CGG TTT CCC CGA GGC CGA CAT CGC 1680  
 D G A R R L H L T W Q S P R M Q S L P E  
 1681 CGA CGG TGC GCG GCG ACT GCA CCT GAC CTG GCA ATC CCC GCG CAT GCA AAG CCT GCC CGA 1740  
 A S R N R L N T V I N T S L P L L A A L  
 1741 GGC CAG CCG CAA CCG CCT CAA CAC CGT CAT CAA CAC CAG CCT GCC GCT GCT GGC GGC ACT 1800  
 H Y D Q L P A L G R L L D L L E A I A R  
 1801 GCA CTA CGA CCA GCT GCC CGC CCT GGG ACG ACT GCT CGA CCT GCT TGA AGC GAT TGC CCG 1860  
 R A A Y L A L L T E Y P Y A L Q R L V R  
 1861 CCG CGC CGC CTA CCT GGC GCT GCT GAC CGA GTA TCC CTA TGC CCT GCA ACG CCT GGT GCG 1920  
 M I G A S G W A A T Y L T R H P L L L D  
 1921 CAT GAT CGG CGC CAG TGG CTG GGC CGC GAC CTA CCT CAC GCG CCA CCC GCT GCT GCT GGA 1980  
 E L L D D R N L K A A S D W A A F A E N  
 1981 TGA ATT GCT GGA TGA CCG CAA CCT CAA GGC CGC CTC CGA CTG GGC CGC CTT CGC CGA GAA 2040  
 C R R Q L A T A E G D T E R Q L D I L R  
 2041 CTG CCG CCG TCA ACT CGC CAC CGC CGA GGG CGA TAC CGA ACG CCA GCT CGA CAT CCT GCG 2100  
 E L H H A E Q F R L L A Q D L E G D L S  
 2101 CGA GTT GCA CCA TGC CGA GCA GTT CCG TCT GCT GGC GCA AGA CCT GGA GGC CGA CCT GTC 2160  
 V E K L A D E L S A L A D V L V Q V T I  
 2161 GGT GGA AAA ACT GGC CGA CGA ACT CTC TGC CCT GGC CGA TGT ACT GGT GCA AGT CAC CAT 2220  
 E A V W Q T V A Q R H R E V P Q F A V I  
 2221 CGA GGC GGT CTG GCA GAC CGT TGC CCA GCG CCA TCG CGA GGT GCC GCA GTT CGC GGT GAT 2280  
 A Y G K L G G K E L G Y A S D L D V V F  
 2281 CGC CTA TGG CAA GCT GGG CGG CAA GGA ACT GGG CTA TGC CTC CGA CCT GGA CGT GGT CTT 2340  
 L Y D D E D Q D A P G L Y A K L A Q R F  
 2341 CCT CTA CGA CGA CGA AGA CCA GGA CGC CCC CGG CCT CTA CGC CAA GCT GGC GCA GCG CTT 2400  
 I T W M T S H T P A G T L F D I D I A L  
 2401 CAT CAC TTG GAT GAC CAG CCA CAC CCC GGC CGG CAC GCT GTT CGA CAT CGA CAT CGC CCT 2460  
 R P D G A S G L L V S P L S S F E K Y Q  
 2461 GCG TCC GGA CGG CGC CAG CGG CCT GTT GGT CTC GCC CTT GTC TTC CTT CGA GAA ATA CCA 2520  
 L N A A W I W E H Q A L T R A R F C A G  
 2521 GCT CAA TGC GGC CTG GAT CTG GGA ACA CCA GGC GCT CAC GCG GGC GCG CTT CTG CGC GGG 2580  
 D T A I G E R F E A L R E R V L R Q P R  
 2581 AGA TAC GGC CAT CGG CGA GCG CTT TGA AGC GCT GCG CGA ACG GGT GCT GCG CCA GCC GCG 2641  
 D P A R L E E E V L S M R R R M R E A H  
 2641 CGA TCC GGC CAG GCT GGA AGA AGA AGT CCT GTC GAT GCG ACG GCG CAT GCG CGA GGC TCA 2700  
 P N R S T M F D L K H D E G G M I D I E  
 2701 TCC CAA CCG CAG CAC CAT GTT CGA CCT CAA ACA CGA TGA GGG CGG CAT GAT CGA CAT CGA 2760  
 F M V Q Y L V L R H A S D H P Q L T G D  
 2761 ATT CAT GGT GCA ATA CTT GGT CCT GCG CCA TGC CAG CGA CCA TCC GCA GCT CAC CGG CGA 2820

Continua

```

      I   G   N   I   A   L   L   K   L   A   G   Q   L   G   L   I   D   A   L   L
2821 CAT CGG CAA TAT CGC CCT GCT CAA GCT GGC CGG ACA GCT GGG GCT GAT CGA TGC GCT GCT 2880
      A   D   E   A   A   N   A   Y   R   L   F   R   K   L   Q   H   Q   I   R   L
2881 GGC CGA GAA GGC CGC CAA TGC CTA TCG CCT CTT CCG CAA GCT GCA GCA CCA GAT CCG CCT 2940
      Q   G   S   E   R   A   H   I   D   A   I   R   V   E   H   E   R   A   C   V
2941 GCA GGG TTC AGA ACG CGC CCA TAT CGA TGC GAT CCG GGT GGA GCA TGA GCG GGC ATG CGT 3000
      I   R   L   W   Q   Q   V   F   G   *
3001 GAT CCG GCT GTG GCA GCA AGT CTT TGG CTG ATC CTG GGC CGG GCC GGC CCC AGG GTG CCG 3060
3061 GTG AAC GAA AAA CGG GAT GCA CCT GCC GGT GTA TCC CGT TTT TCA TGT CCA GCG CTG AGC 3120
3121 CAA CAG CGA GGT CGC C

```

A sequência de bases do gene *glnE* é mostrada em azul. Os códons de início (ATG) e término (TGA) de tradução estão sublinhados. Acima da sequência de bases está a sequência traduzida em aminoácidos obtida com o programa Strider. A montante ao códon de início de tradução há um provável sítio de ligação para ribossomo (em vermelho), duas possíveis regiões promotoras dependente do fator  $\sigma^{54}$  (em verde – promotores 1 e 2) e uma possível região promotora dependente do fator  $\sigma^{70}$  (em roxo – promotor 3). A jusante ao códon de parada de transcrição (TGA) há uma sequência palindrômica (em negrito) que pode atuar com terminador de transcrição.

**FIGURA 12: Identificação de promotores**

<b>FATOR <math>\sigma^{54}</math>:</b>	
Seqüência consenso:	<b>CTGGYAYRNNNNTTGCA</b>
Promotor 1:	<b>GCGGTGACATGATCGCA</b>
Promotor 2:	<b>CTGGGTGCAAAATGGCC</b>
 <b>FATOR <math>\sigma^{70}</math>:</b>	
Seqüência consenso:	<b>TTGACA_____N<sub>15-17</sub>_____TATAAT</b>
Promotor 3:	<b>TAAATTGAAAAATGAAGGATAGAAT</b>

Comparação entre as seqüências consenso e aquelas localizadas a montante ao gene *glnE* para os promotores dependentes de  $\sigma^{70}$  e  $\sigma^{54}$ . Em negrito estão as bases conservadas. Na seqüência consenso de  $\sigma^{54}$  “Y” representa bases pirimídicas e “R” representa bases púricas.



**FIGURA 13: Comparação da sequência de aminoácidos da proteína GlnE de *H. seropedicae* com proteínas GlnE de *Ralstonia solanacearum*, *Burkholderia fungorum*, *Cromobacterium violaceum*, *Escherichia coli* e *Azotobacter vinelandi*.**

CLUSTAL W (1.82) multiple sequence alignment

```

R.solanacearum      -----MTAQAPLSVAPFLQ 14
B.fungorum          -----MLLRIEQRLS 10
H.seropedicae       -----MAAGFPSTLCCGAFQ 16
C.violaceum         -----
E.coli              -----MKPLS-----SP---LQQY-----WQTVVERLP 20
A.vinelandi         MGGCGIVVSGRRVIRLSCSFITQWRFPSEDFCKITPRVGLNRFGKSGRYQDTSMSLP 60

R.solanacearum      TPACPSEAPGAVLPFSLN-----YSRVQRLA 41
B.fungorum          SNERARPLKTDATLLSSS-----YSHYAARA 37
H.seropedicae       TLMPVFSPAHHRSIVEAS-----SSRYTRWT 43
C.violaceum         ---MPANTSAAIARSCE-----FSQYLHRL 23
E.coli              ---EPLEESLSAQAKS-----VLTFSDVFQDSV 46
A.vinelandi         TLIDLPPALETLGNQALQRFGEAAGTLGDSPATPFADWPPARRAAAFARVCANSQGVFLERA 120
                        * :

R.solanacearum      QARA-GWAERVQAAAAG-PISVAWL SARFGELEFAAAAEP-----EQALKRAIRLLRNEV 94
B.fungorum          AARP-QLVAQVAALASA-PLTRERIDARFDALCAEAAGASGAPLSEDAIKRAIRQLRTEV 95
H.seropedicae       QADP-ARPALVAAAAAGALDRAFLQQRLEHDLASGLSLP-----AAMRRLRLNLV 92
C.violaceum         TARPERQRLEQSLHHPFSLQAFADWPALDAPEALAP-----AIRRLRQAV 72
E.coli              IAHPEWLTELESQP--PQADEWQHYYAAILQEAICNVSDA-----GLMRELRLFRRI 97
A.vinelandi         LRDPAMLLQLAGDGLSLERSLAPDELCCQLAALLADCAQEE-----ELGRRLRFRHRQ 173
                        : * :

R.solanacearum      FAALAERDLSCAATLDEVTGAMTDAEFVAVRAIAVIGQELGALHGQPVGSSGEVQELV 154
B.fungorum          FCAVMERDLAGEADVAEVTGAMTDLAETTIQRALAVLSAELEVLYGEPRG-PHGERLSLG 154
H.seropedicae       ICTLITRDLGRGDLDEVVTMTGFADFVQTHLAALMQEQTALYGLPIGEESGRPQEMI 152
C.violaceum         VARLICRDLESLATLDEVVSTISQLAEFAVRQALACAAASLP-QYGRPIGEDSGEPQQLI 131
E.coli              MVRIAWAQTALVTEESILQQLSYLAETLIVAARDWLYDACCWGTGTPCN-AQGEAQPLL 156
A.vinelandi         QVRIVWRDLMRQADLAETCRDLSELADACIDLACDWLYERLRARFGTPLGRSGRPQRLV 233
                        : : : : : * * * * :

R.solanacearum      VVGMGKLGRELNVSSDIDLIFLYDEEGDT-QGGPR-PLSNHEYFTKLGRRLINALADVT 212
B.fungorum          VVGMGKLGRELNVSSDIDLIFYEEDGET-AGGHRSPATQEFFTRLGKRLIGALAEVT 213
H.seropedicae       VVGMGKLGRELNVSSDIDLIFVYPEDGDTAEAGQKSLSNHEFFVRLGKRLIGALAEIT 212
C.violaceum         VVGMGKLGRELNVSSDIDLIFYPEGET--DGAR-KISNHEYFTQVGKRLIALLNAT 188
E.coli              ILMGKLGRELNVSSDIDLIFAWPEHGCT--QGGRELDNAQFFTRMGQRLIKVLDQPT 214
A.vinelandi         VVGMGKLGRELNVSSDIDLIFYPEGET--DGERRLDNQEFFVRLGQKLIGALDAIT 291
                        : : : : : * * * * :

R.solanacearum      EDGYVFRVDMRLRPNGDAGPLACSLGMLEEYFVVGQGREWERYAWIKGRVITDPGSPHAA 272
B.fungorum          ADGYVFRVDMRLRPNGDAGPLVCSLGMLEEYFVVGQGREWERYAWIKGRVISEGTSDAAR 273
H.seropedicae       EDGFTFRVDMALRPNGNSGPLVASFNMVVEYLVRQGREWERYAWTKARALTGTPEDIAT 271
C.violaceum         ADGQVFRVDMRLRPYGDSPGLVMSFAALENYLLSQGREWERYAWIKAKAITGDADGLAQ 247
E.coli              QDGFVFRVDMRLRPFGESGPLVLSFAALEDYQEQGRDWERYAMVKARIMGDSEGVYAN 273
A.vinelandi         VDQGAFRVDMRLRPYSSSGLVHSFGAMEQYYQDQGRDWERYAMIKARVVG-GDREAGR 349
                        * * : * * * * * : * * : * * :

R.solanacearum      VIQQLERVTPFVFRRYLDYGVIAAIRALHAQIRAEAAKRSNAANRQHGGHDSQVQAHAP 332
B.fungorum          LQKQLDAIVTPFVFRRYLDYGVISAIRALHLQIRQEAQRASMR-----PDKAD 322
H.seropedicae       ---LEAISRPFIFRRYLDYGSIDALRSMHGQIRAEVQRQEAALH-----PDRSN 316
C.violaceum         ---LVRPFVFRRYLDYGVIAAIRALHAQIRAEVARR-----DMAD 284
E.coli              ---ELRAMLRPFVFRRYLDYGVISAIRALRDMKQIRQEVRRK-----GLTD 314
A.vinelandi         ---QLLELLRPFVFRRYLDYGVIAAIRALRDMKQIRQEVRRK-----GMAD 390
                        : * * : * * : * * : * * :

```

Continua



R.solanacearum	NIKLGRGGIREIEFVAQVFQLIRGGQDFGLRIRPTLEVLRAAAERGLIGADAVARLTDAY	392
B.fungorum	DIKLGRGGIREIEFSAQVFQLIRGGQDAGFRVRPTLAVLRHAATHGLIDTSVCVKLSQAY	382
H.seropedicae	NVKLGRGGIREIEFTSQVFQLIRGGRDALDRSTRITLRTLAAKDLLAPEVVEQLLHAY	376
C.violaceum	NIKLPGGIREAEFIAQVFQLIRGGRDRTLQLRGTRATLERLAALRLLEPAVAELQASY	344
E.coli	NIKLGAGGIREIEFIVQVFQLIRGGREPSLQSRSLPRLTSAIAELHLLSENDAEQLRVAY	374
A.vinelandi	NVKLGAGGIREVEFVAQVFQLIHGGDRSLQRRSLLEVLATLEGQGYLPPAAVAELREGY	450
	: : * * * * * * * * * * : : : * : *	
R.solanacearum	RFRLRQLEHRLQYVDDAQTHNLPGAPEDQLRIARMMGFADYAAALVAQLERYQDEVAAQQFEQ	452
B.fungorum	RFRLRELEHRLQYRNDAAQTHAMPVDPEERVALAHAMGCDYAAALMAKLDAREFVEQQFDQ	442
H.seropedicae	TFRLRNLEHRLQYLEDAQTHTLVPNPDDLVLVARMMGYEPAALLHELERQRAIVAAQFDA	436
C.violaceum	AFRLRNLEHRLQYLDQQQTQLPEAPETRQKIAASMGHADWPAFLDALNEVRRKVSRRHEQ	404
E.coli	LFLRLLENLQSIINDEQTTLPDSELNRLARLAWAMDFADWPQLTGALTAMHTNVRVFNFE	434
A.vinelandi	EFRLRYTEHALQAIADRQTFMLPNELDRLARVACMLGCAGWEAFGLRLAHWRHEFRQ	510
	* * * * * : * * * * : * : * : *	
R.solanacearum	TFSDKQDNQPPCAA-----IWHADLDDERTES-ARALQLEELGYADA--DSVLRLR	501
B.fungorum	IFADKVSGRDGGCAGPEDGAAAWVWSSALADDSADDV-LQARLVELGVAE--GELLARLR	499
H.seropedicae	IFADKQSEADNDGP-----SVSVSDNDNLEG-LADALRLVGFPEADIADGARLRH	485
C.violaceum	VFILPSEDSASHPL-----SELWLDVAEQS-PETRLAELGYADP--AAVARQLT	450
E.coli	LIGD--DESETQES-----LSEQWRELWQDALQEDDTTPVLAHLSDDR--KQVRLTLIA	485
A.vinelandi	VIAVPGEDEASEGRP----VDAQWLPWLWENGDEAAVCRQLVEAGFADP--QVVRQLRA	563
	: : : : * : :	
R.solanacearum	ASRHSPPRYALSEVSRQRFDLINRALDHAARQTDAD---VTIARFLDFFDATSRSSYL	558
B.fungorum	AVWQSSRYAGLAERSRQRFDIVAQRALEAARTLEPAERRGDTVARFFDLLEAVSRRGAYL	559
H.seropedicae	LTWQSPRMQSLPEASRNRLNTVINTSLPLLAALHYDQLP--ALGRLLDLLEATARRAYL	543
C.violaceum	GLAQSQRYLQMPLAGRQQLDALMPALIEVAARFPNAD---DTLSRIIGLMEATSRASYL	507
E.coli	DFRKELDKRTIGPRGRQVLDHLMFHLLSDVCAREDA---VTLSRITALLVGIVRTTYL	542
A.vinelandi	ALRGGSQVRGMQRLGRERLDLAFIPQLLARVVEQDAPD---LVLERVLPLVEAVARRSAYL	620
	: : * : : : : : * : : *	
R.solanacearum	SLLSEYPQAMARVAHTLHASRWAADYLTRHPQLLDELDDTEALSAAPDWQGFRRVRERL	618
B.fungorum	ALLTEYPQALHHVLSVLGGSRWAAGYLIRHPQLLDELDDDEAMDSPFDWPEFKRTLRLR	619
H.seropedicae	ALLTEYPYALQRLVRMIGASGWAATYLTRHPQLLDELDDDRNLKAASDWAAPFENCRRQL	603
C.violaceum	ALLTEYPQTLQRLASLYSSSAVSAAYLSRHPILLDELDDARMLYAAPDWPLLAQLETQL	567
E.coli	ELLSEFPAAALKHLISLCAASPMVASAQRLARYPLLLDELDPNTLYQPTATDAYRDLRQYL	602
A.vinelandi	VLLTENPGALRRLSLCAASPMIAEQIARFPLLLDQLLNEDRLYSPQEAELVAELRELL	680
	* * : * : : : * : : * * * * * : :	
R.solanacearum	RAAG-DHVEQMMDILRQEHHAETFRILLQDLQGMILTVEPISDRLSDLADAVLDLTLETW	677
B.fungorum	AAA--DGVEQQMDLIRHAHQAEVFRILLIDLAKGLSVEHVSDRLSELADAVLDVTLEAVW	677
H.seropedicae	ATAE-GDTERQLDILRELHHAEQFRLLAQDLEGLDSVEKLADELSALADVLVQVTIEAVW	662
C.violaceum	AQAD-GDVEAKMDALRHFAQHQAQFRLVAQDLAAGMTLEALSDELSRLADLVLAABVRHAW	626
E.coli	LRVPEDDEEQQLEALRQFKQAQLLRIAAADIAGTLPMVKVSDHLTWLAEAMIDAVVQQAW	662
A.vinelandi	LRIPEDDLEQQMEALRHFKLAHLRLVAASEIAGTLPLMKVSDYLTWLAETLDQALRLAW	740
	. * : : * * : * : * : : * : : * * : : * : *	
R.solanacearum	QQVSTRH-----REVPRFAVVAYGRLGGKELGYASDLDLIFLYD-----DDD	719
B.fungorum	RQLPKRH-----RDVPRFAVIAYGKLGGKELGYASDLDVIFLYD-----DTD	719
H.seropedicae	QTVAQRH-----REVPRFAVIAYGKLGGKELGYASDLDVIFLYD-----DED	704
C.violaceum	RDIPSRH-----CETPRFAVIYGYKLGKELGYASDLDLIFLYD-----DEH	668
E.coli	VQMVARYGKPNHLN-EREGRGFVVVGYGKLGWELGYSSDLDLIFLHDCPMDAMTDGERE	721
A.vinelandi	RQMVVRHGNPRRSRGSPCELDLFLVGYGKVGIGELHGSGDLDLFIHGNPQAEITDGAKP	800
	: * : * : * * * * * * * * * * * * * * *	
R.solanacearum	ERAPDEVYAAARKLITWLTSHTAAGMLFDVDTRLRPNGAAGLMVTHFEAFRRYQMREGD-	778
B.fungorum	EAAAEVYSTYTRRLITWLTTATGAGTLFDVDLRLRPNGESGLLVTDLDAFRRYQLREGDA	779
H.seropedicae	QDAPGLYAKLAQRFITWMTSHTPAGTLFDIDIALRPDGASGLLVSPLSSEFKYQL----	759
C.violaceum	PDAPDLYSRLARKLSTWLTSAAGVLYDIDLRLRPNGSSGLLVSSISAFRRYQ-----	723
E.coli	IDGRQFYLRLAQRIMHLFSTRTSSGILYEVDARLRPSGAAGMLVTSAAEFADYQK-----	776
A.vinelandi	IDGAQFFTRLGQRIVHLLTAQTPSGQLYEVDMLRLRPSGNAGLLVSSLGAFRRYQ-----	855
	. * * * * * * * * * * * * * * * * * *	
R.solanacearum	-NAAWVWEHQALTRARFCAGDPEIGARFEALRIAVLRQPREAGPLRDEIAAMRERVLEGH	837
B.fungorum	ANTAWVWEHQALSRARYCAGDAGIKAFEAIREQVLTTPREAAPLAKEIVEMRERVVAGH	839
H.seropedicae	-NAAWIWEHQALTRARFCAGDTAIGERFEALRERVLRQPRDPARLEEVLSMRMRMREAH	818
C.violaceum	-NQAWVWEHQALTRARFVAGDAGTGSQFAERHAILTLERDPAKLREVMAMRQRMMLDSH	782
E.coli	-NEAWTWEHQALVRARVVYQDPQLTAHFDAVREIMTLPREGKTLQTEVREMREKMRHL	835
A.vinelandi	-HEAWTWEHQALVRARALVGSPLAAAFAEVRAQVLRGRPRDLAVLRREVGDMRAKMRSL	914
	. * * * * * * * * * * * * * * * * * *	



**FIGURA 14: Domínios estruturais da proteína GlnE**a) *Herbaspirillum seropedicae*b) *Ralstonia solanacearum*c) *Escherichia coli*

A identificação dos domínios estruturais foi realizada pelo programa SMART. Tanto a proteína GlnE de *H. seropedicae* como as de *R. solanacearum* e *E. coli* apresentaram dois domínios estruturais: um na região N-terminal correspondente ao sítio de desadenilação e outro na região C-terminal correspondente ao sítio de adenilação. RPT indica domínios repetidos (do inglês *RePeaT*). PFAM indica famílias de proteínas (do inglês *Protein FAMilies*).



### 4.3 Mutagênese do gene *glnE*

O plasmídeo H11 foi utilizado para mutagenizar o gene *glnE* através da inserção de um cassete de canamicina na região intermediária do gene.

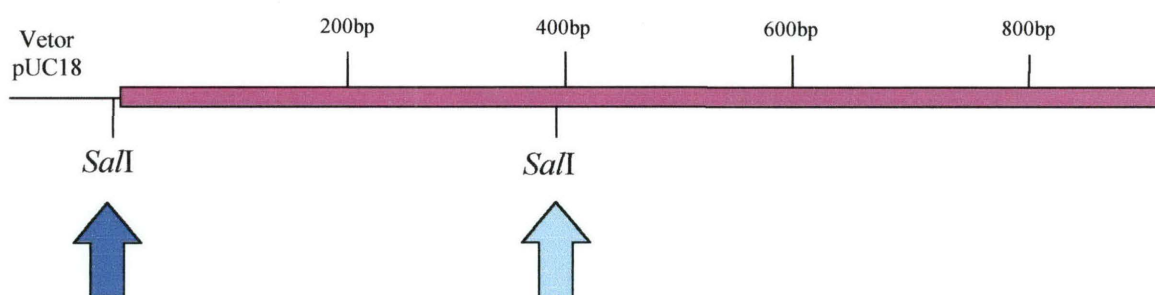
A sequência de DNA obtida com o iniciador universal foi utilizada para a elaboração de um mapa de restrição eletrônico da região inicial do gene *glnE* com o propósito de se identificar sítios únicos para enzimas de restrição. Um sítio para a enzima *SalI* foi localizado e, apesar de existir um sítio para essa enzima no vetor pUC18, ele foi escolhido como local para a inserção do cassete de canamicina (figura 15). Nessa região o gene *glnE* apresenta sequências a montante e a jusante longas o suficiente para que ocorra pareamento no momento da recombinação homóloga. O sítio *SalI* também pode ser utilizado para liberar o cassete de canamicina intacto a partir do plasmídeo pUC4K (anexo1).

Para confirmar se o sítio de restrição para *SalI* era único no inserto, o DNA do clone foi tratado com a enzima *SalI*. Na foto da eletroforese (figura 16) pode-se visualizar duas bandas decorrentes da atividade da enzima *SalI*, uma banda com aproximadamente 4.000 pares de bases e outra com aproximadamente 400 pares de bases. Esse padrão corresponde a apenas dois sítios de restrição e, como já era esperado, um deles está no vetor enquanto o outro está presente dentro da sequência do gene *glnE* no plasmídeo H11. A banda menor corresponde ao fragmento de DNA existente entre o sítio *SalI* do vetor e do inserto. A banda maior corresponde ao vetor pUC18 (cerca de 2.700 bases) com o restante do inserto (aproximadamente 1.300 bases).

A restrição total do plasmídeo H11 com a enzima *SalI* libera um pequeno fragmento de aproximadamente 400 pares de bases. Esse fragmento contém a parte inicial do gene *glnE* e uma pequena parte do vetor. Para se inserir o cassete no local desejado foi necessário se obter um fragmento derivado de restrição parcial, na qual apenas o sítio que fica dentro do gene *glnE* estivesse aberto. Com esse fim foi feito um sistema de restrição (item 3.8), do qual alíquotas de 10uL foram

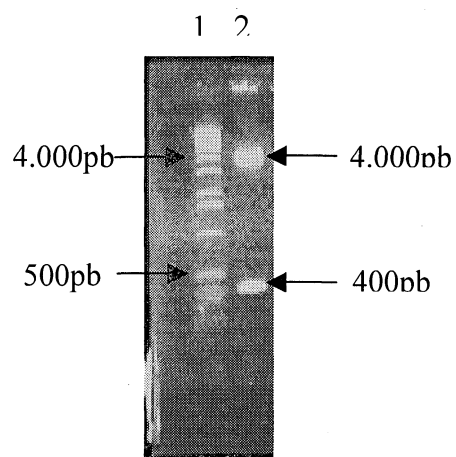
retiradas nos tempos indicados (figura 17). O fragmento de DNA utilizado para a mutagênese foi aquele produzido durante a incubação de 15 minutos, pois nos tempos posteriores já havia sido formada uma banda mais intensa de aproximadamente 4kb correspondente ao perfil de restrição total (figura16). O DNA incubado durante 15 minutos poderia se apresentar de 3 formas: 1) apenas o sítio no gene estar cortado (nosso objetivo), 2) apenas o sítio do vetor estar cortado e 3) ambos os sítios estarem cortados.

**FIGURA 15: Sítios de restrição para a enzima *SaI* existentes no plasmídeo H11.**



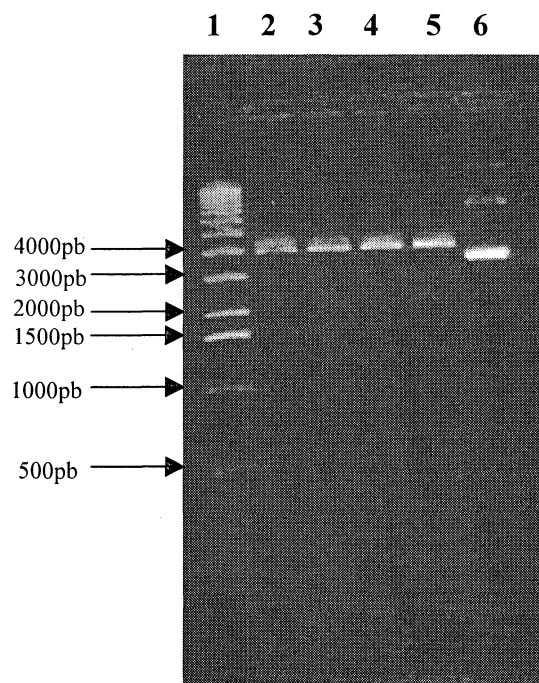
A análise do mapa eletrônico de restrição indicou a existência um sítio *SaI* aproximadamente na base 390 da sequência. Esse sítio (seta verde) foi utilizado como alvo para a inserção do cassete de canamicina. Também está indicado o sítio *SaI* existente no vetor (seta azul). O mapa eletrônico de restrição foi elaborado com o auxílio do programa disponível no sítio ([www.expasy.org](http://www.expasy.org)).

**FIGURA 16: Restrição do plasmídeo H11 com a enzima *SalI***



1) Marcador de peso molecular 1 kb Ladder; 2) Plasmídeo H11 tratado com *SalI*.

**FIGURA 17: Curva de restrição do plasmídeo H11 com a enzima *SalI*.**



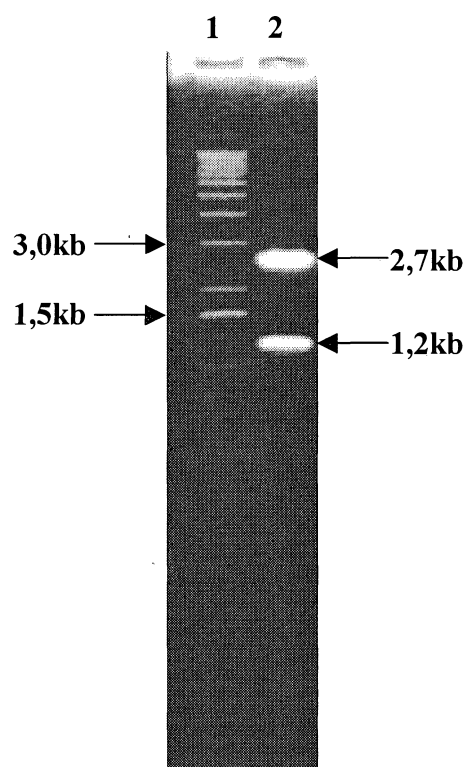
1) marcador de peso molecular 1 kb Ladder; 2) alíquota coletada em 15 minutos; 3) alíquota coletada em 30 minutos; 4) alíquota coletada em 45 minutos; 5) alíquota coletada em 60 minutos; 6) plasmídeo H11 fechado. O plasmídeo H11 foi tratado com a enzima *SalI*. Alíquotas de 10uL foram coletadas do sistema de restrição nos tempos de 15, 30, 45 e 60 minutos.



O plasmídeo pUC4K também foi tratado com a enzima de restrição *SalI* visando liberar o gene de resistência à canamicina do vetor. A foto da eletroforese (figura 18) mostra duas bandas, uma de aproximadamente 1,2kb, correspondente ao cassete de canamicina, e outra de aproximadamente 2,7kb, correspondente ao vetor. O DNA digerido com a enzima *SalI* foi então aplicado em gel de baixo ponto de fusão e a banda de 1,2kb foi extraída (item 3.7.1).

O plasmídeo H11 digerido parcialmente com a enzima *SalI* foi misturado em um sistema de ligação com o DNA extraído do gel de baixo ponto de fusão correspondente ao gene de resistência a canamicina (banda de 1,2kb). A construção resultante desse sistema, assim como a restrição parcial do plasmídeo H11, poderia se apresentar de 3 formas distintas: 1) o cassete de canamicina poderia entrar no sítio de restrição do gene *glnE* (figura 19a); 2) o cassete de canamicina poderia entrar no sítio de restrição presente no vetor (figura 19b); ou 3) o pequeno fragmento de DNA entre os dois sítios *SalI* poderia ser perdido e o cassete de canamicina entraria no seu lugar (figura 19c). Dessas três possibilidades, apenas a primeira era de interesse.

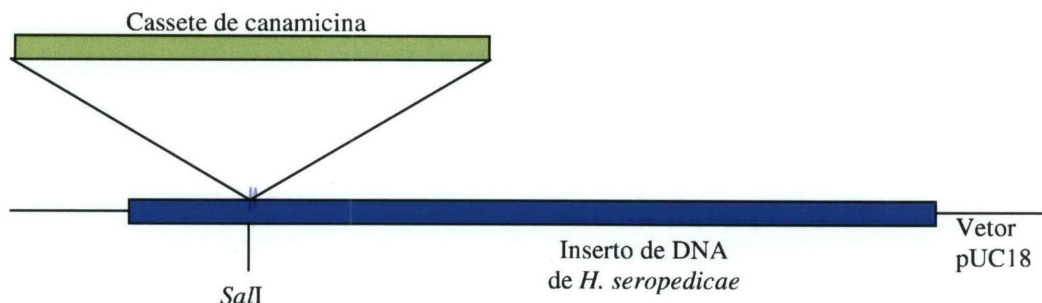
**FIGURA 18: Restrição do plasmídeo pUC4K com a enzima *SalI***



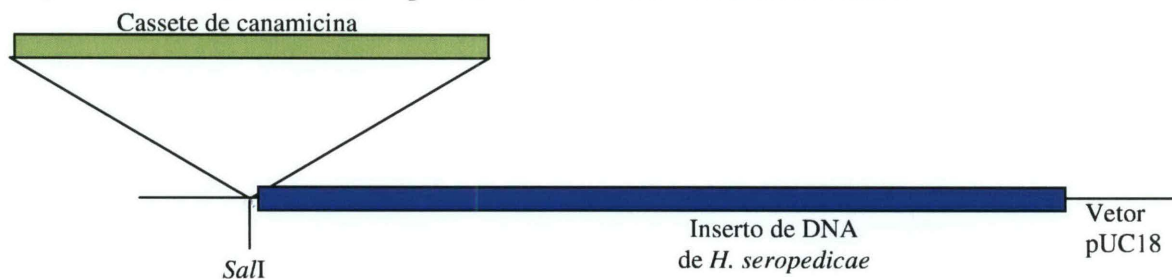
Identificação das colunas: 1) marcador de peso molecular 1 kb Ladder; 2) pUC4K tratado com *SalI*. A restrição do pUC4K com *SalI* originou duas bandas. Uma menor com cerca de 1,2kb, correspondente ao gene de resistência à canamicina, e outra banda maior com aproximadamente com aproximadamente 2,7kb correspondente ao vetor.

**FIGURA 19: Possibilidades de inserção do cassete de canamicina no plasmídeo H11**

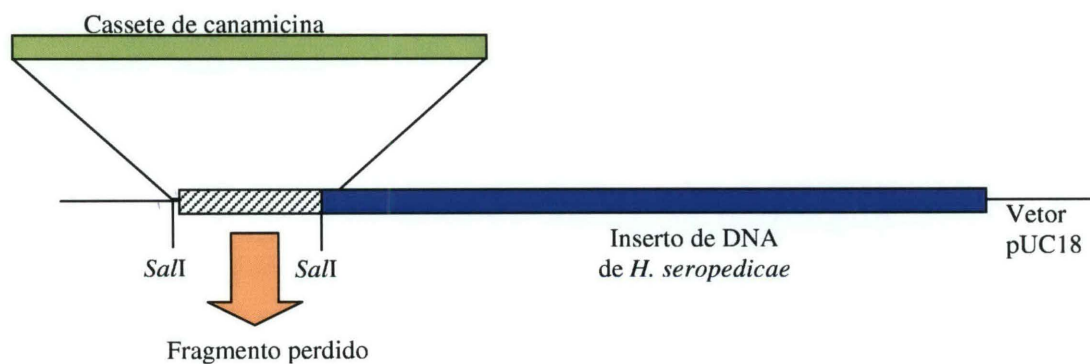
a) O cassete de canamicina poderia entrar no sítio *SalI* do inserto:



b) O cassete de canamicina poderia entrar no sítio *SalI* do vetor:



c) O fragmento de 400 pares de bases entre os sítios *SalI* poderia ter sido perdido e o cassete de canamicina entraria no seu lugar:



O sistema de ligação foi eletroporado na estirpe DH10B de *E. coli* (item 3.6). As bactérias transformadas foram plaqueadas em meio sólido contendo o antibiótico canamicina para seleção de clones que continham o gene de resistência à canamicina. Como resultado foram isoladas 41 colônias que foram coletadas e transferidas para uma placa mestra. Os plasmídeos foram purificados (item 3.3.1) e passaram a compor a série pHF.

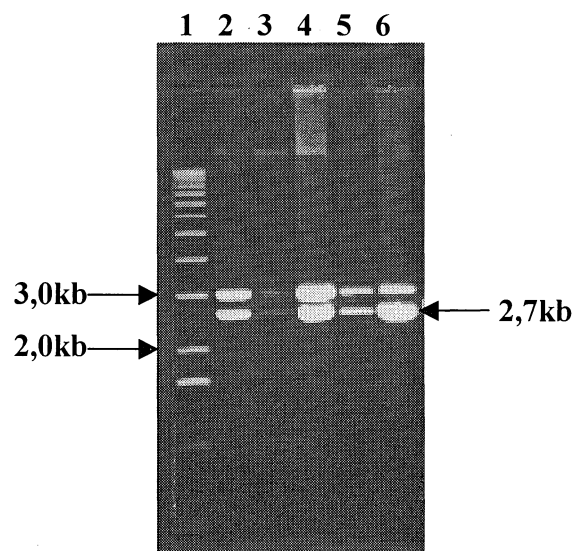
A identificação dos clones pHF que possuíam o plasmídeo com o gene *glnE* mutado foi feita através de análise do padrão de restrição do DNA (item 3.3). Nesse procedimento foram usadas as enzimas de restrição *XbaI* e *SacI* com a intenção de liberar o inserto do vetor, já que essas enzimas cortam o vetor pUC18 mas não cortam o inserto. Para cada uma das três possibilidades de inserção do cassette de canamicina nos sítios *SalI* do plasmídeo H11 (figura 19a, 19b e 19c), havia um padrão de restrição diferente:

- 1) Se o cassette de canamicina entrasse no sítio *SalI* do inserto, a restrição com *XbaI* e *SacI* apresentaria duas bandas, uma correspondente ao vetor pUC18 com aproximadamente 2.700 pares de bases e outra banda com aproximadamente 3.000 pares de bases correspondentes ao inserto mais o cassette de canamicina.
- 2) Se o cassette de canamicina entrasse no sítio *SalI* do vetor, a restrição apresentaria também duas bandas, mas de tamanhos diferentes das bandas anteriores. Uma banda, correspondente apenas ao inserto, teria aproximadamente 1,7KB enquanto a outra banda, formada pelo vetor pUC18 mais o cassette de canamicina, teria em torno de 3,9kb.
- 3) Se o fragmento de 400 pares de bases entre os sítios *SalI* do vetor e do inserto fosse perdido e o cassette de canamicina entrasse em seu lugar, apenas uma banda de aproximadamente 5,2kb poderia ser vista, já que junto com o fragmento de 400kb também seria perdido o sítio para a enzima *SacI*.

Dos 41 clones analisados, cinco apresentaram o padrão de restrição descrito na primeira possibilidade: pHF16, pHF30, pHF33, pHF36 e pHF38. Na figura 20 pode-se observar o padrão de restrição destes clones.

O clone pHF16, contendo o plasmídeo com o gene *glnE* mutado, foi selecionado para a obtenção de um mutante *glnE* cromossomal de *H. seropedicae*. Este plasmídeo foi introduzido na estirpe selvagem SMR1 de *H. seropedicae* por eletroporação (item 3.6). As células mutantes foram selecionadas em meio NFbHPN sólido contendo os antibióticos estreptomicina (marca da estirpe selvagem SMR1) e canamicina. O antibiótico canamicina foi usado em uma alta concentração de 0,5mg/mL, superior àquela utilizada para *E. coli* (0,05mg/mL), pois *H. seropedicae* possui uma resistência natural à canamicina. Sessenta e oito colônias foram isoladas e transferidas para uma placa mestra. Essa nova estirpe foi denominada HEL1 a HEL68.

**FIGURA 20: Confirmação da inserção do cassete de canamicina no plasmídeo H11**



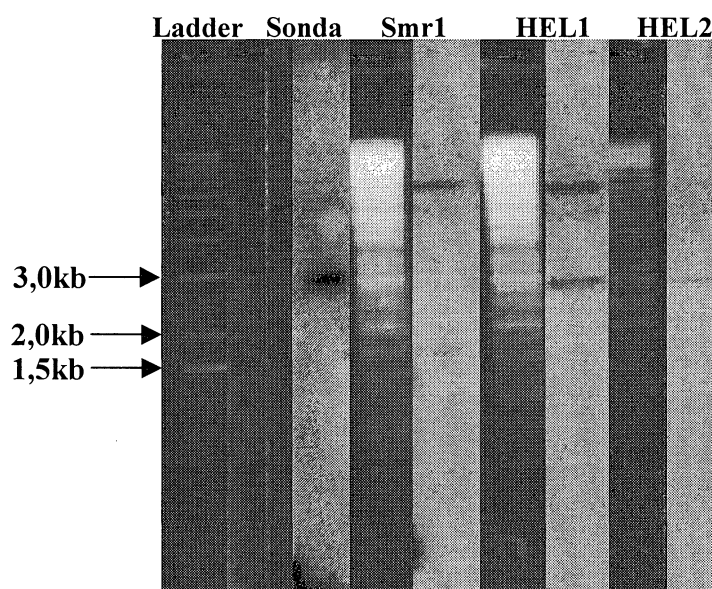
1) marcador de peso molecular 1 kb Ladder; 2) pHF16; 3) pHF30; 4) pHF33; 5) pHF36; 6) pHF38. A restrição do DNA dos clones pHF16; pHF30; pHF33; pHF36 e pHF38 com as enzimas *Xba*I e *Sac*I originou duas bandas: uma com 2,7kb correspondente ao vetor pUC18 e outra com aproximadamente 3kb correspondendo ao inserto de DNA de *H. seropedicae* mais o cassete de canamicina. Esse padrão confirma a mutação no plasmídeo H11.

A confirmação da mutação no gene *glnE* foi feita através de hibridização. Para tanto, o DNA genômico da estirpe selvagem SMR1 e das estirpes mutantes HEL1 a HEL5 foi extraído (item 3.3.2) e submetido à digestão total com a enzima de *EcoRI* gerando um padrão de bandas. O DNA digerido foi hibridizado com uma sonda específica que, nesse caso, tratava-se do plasmídeo pHF16 também digerido com *EcoRI*. A mutação em *glnE* foi confirmada pela presença de uma banda eletroforética de aproximadamente 3kb nas estirpes HEL1 a HEL5, correspondente ao gene *glnE* mais o cassete de canamicina, que está ausente na estirpe selvagem SMR1

O perfil de hibridização permitiu observar que a mutagênese das estirpes HEL1 e HEL2 ocorreu por dupla recombinação. A estirpe selvagem SMR1 apresenta uma banda hibridizada, correspondente ao gene *glnE* de aproximadamente 1,8kb. Nas estirpes mutantes HEL1 e HEL2 essa banda teve um acréscimo de 1,2kb em relação à estirpe selvagem SMR1. Esse acréscimo corresponde ao cassete de canamicina (figura 21). As outras estirpes mutantes apresentaram um perfil de bandas distinto e indicativo de que tenha ocorrido apenas simples recombinação (dados não mostrados).



**Figura 21: Confirmação da construção de estirpes mutantes para o gene *glnE* em *H. seropedicae*.**



As estirpes mutantes HEL1 e HEL2 apresentam uma banda com aproximadamente 3kb que está ausente na estirpe selvagem. Essa banda corresponde ao gene *glnE* mais o cassete de canamicina. A sonda utilizada para hibridizar o DNA foi o plasmídeo pHF16 tratado com a enzima *EcoRI*.

#### 4.4 Caracterização fisiológica dos mutantes

A capacidade de fixação de nitrogênio dos mutantes foi avaliada (item 3.10). A estirpe mutante HEL1 apresentou uma redução de aproximadamente 29,5% na capacidade de fixação de nitrogênio em relação à estirpe selvagem SMR1 (tabela 2). Espera-se que com o gene *glnE* inativo, a glutamina sintetase sintetizada permaneça na sua forma desadenililada, ou seja, na forma ativa, mesmo em um meio com excesso de amônio. A regulação da atividade de GS por adenililação é gradual, isto é, a atividade diminui de acordo com o número de subunidades de GS que estão adenililadas ficando totalmente inativa quando as doze subunidades estão adenililadas. Desse modo a estirpe selvagem SMR1 pode vir a apresentar uma atividade de GS reduzida em relação ao mutante devido a adenililação de somente algumas subunidades. Sugerimos que no mutante HEL1, a manutenção de GS na forma ativa possa acarretar um acúmulo maior de glutamina intracelular. Uma vez que altos níveis de glutamina intracelular indicam uma condição de excesso de nitrogênio e estimulam atividade removedora de UMP de GlnD, GlnD desudirililada converte PII-UMP em PII livre. Em sua forma livre, PII se liga a NtrB estimulando sua atividade fosfatase que converte NtrC-P em NtrC. Dessa forma NtrC não pode ativar a transcrição do gene *nifA* o que acaba por reduzir a transcrição dos genes estruturais do complexo da nitrogenase e, conseqüentemente, a fixação de nitrogênio no mutante *glnE* (tabela 2).

**TABELA 2: Efeito da mutação em *glnE* sobre a atividade da nitrogenase em *H. seropedicae***

<b>Estirpe</b>	<b>Condições</b>	<b>Tempo de incubação</b>	<b>Atividade específica (<math>\eta</math> mol de etileno formado/proteína.min)</b>
SMR1	20mM $\text{NH}_4^+$	60 minutos	$0,09 \pm 0,03$
SMR1	0,5mM glu	60 minutos	$15,94 \pm 1,81$
HEL1	20mM $\text{NH}_4^+$	60 minutos	$0,08 \pm 0,01$
HEL1	0,5mM glu	60 minutos	$11,24 \pm 1,16$

Os dados se referem a 4 repetições.

## 5. CONCLUSÃO

- O plasmídeo HS03-FP-00-000-014-H11, contendo um fragmento do gene *glnE* de *H. seropedicae*, foi identificado no banco de dados do Programa GENOPAR;
- A transcrição do gene *glnE* de *H. seropedicae* provavelmente dependente de dois possíveis promotores dependentes do fator sigma 54 e um dependente do fator sigma 70;
- Através do experimento de mutagênese sítio dirigida foram obtidas estirpes mutantes de *H. seropedicae* para o gene *glnE* denominadas HEL1 e HEL2;
- Aparentemente, a capacidade de fixação de nitrogênio da estirpe mutante foi afetada pela mutagênese do gene *glnE*.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALTSCHUL, S. F.; MADDEN, T. L.; SCHAFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic. Acids Res., v.25, p.3389-3402, 1997.
- ANDERSON, W.B.; STADTMAN, E.R. Glutamine synthetase deadenylation: A phosphorolytic reaction yielding ADP as nucleotide. Biochem. Biophys. Res. Commun., v.41, p.704-709, 1970.
- ARCONDÉGUY, T.; JACK, RACHAEL.; MERRICK, M. PII signal transduction proteins, pivotal players in microbial nitrogen control. Microbiology and Molecular Biology Reviews, p.80-105, 2001.
- AUSUBEL, F.M. Regulation of nitrogen fixation genes. Cell, v.37, p.5-6, 1984.
- BALDANI, J.I.; BALDANI, V.L.D.; SAMPAIO, M. J. A. M.; DOBEREINER, J. A fourth *Azospirillum* species from cereal roots. An. Acad. Bras. Cienc., v.56, p.365, 1984.
- BALDANI, J.I.; BALDANI, V.L.D.; SELDIN, L.; DÖBEREINER, J. Characterization of *Herbaspirillum seropedicae* gen., nov, sp., nov., a new root-associated nitrogen-fixing bacterium. Intl. J. Sys. Bact., v.36, p.86-93, 1986.
- BENELLI, E.M.; SOUZA, E.M.; FUNAYAMA, S.; RIGO, L.U.; PEDROSA, F.O. Evidence for two possible glnB type genes in *Herbaspirillum seropedicae*. J. Bacteriol, Washington, v.179, p.4623-4626, 1997.
- BURNS, R.C.; HARDY, K.W. Nitrogen Fixation in bacteria and higher plants. Berlim: Spring-Verlag, 1975.
- CABAN, C.E.; GINSBURG, A. Glutamine synthetase adenylyl transferase from *Escherichia coli*: purification and physical and chemical properties. Biochemistry, v.15, p.1569-1580, 1976.

- DIANESE, J. C.O.; DÖBEREINER, J., DOS SANTOS, L. T. Membrane protein patterns of three *Azospirillum* species and *Herbaspirillum seropedicae* strain Z78. Can. J. Microbiol., v.61(2), p.3168-75, 1989.
- DILWORTH, M.J. Acetylene reduction by nitrogen-fixing preparations from *Clostridium pasteurianum*. Bioch. Biophys Acta, v.127, p.285-294, 1966.
- DÖBEREINER, J. History and new perspectives of diazotrophs in association with non leguminous plants. Symbiosis, Rehovot, v.13, p.1-13, 1992a.
- DOUGLAS, S. E. DNA Strider: a Macintosh program for handling protein and nucleic acid sequences. Methods Mol Biol, v.25, p.181-94, 1994.
- EPSTEIN, E. Nutrição mineral das plantas princípios e perspectivas. São Paulo: Universidade de São Paulo, p.341, 1975.
- EADY, R.R. Enzymology in free-living diazotrophs. In: BRPUGHTON, W.J.; PULHER, S. eds, Nitrogen Fixation, v.4. Oxford: Clarendon Press, p.1-49, 1986.
- FALK, E. C.; JOHNSON, J. L.; BALDANI, V. L. D.; DOBEREINER, J.; KRIEG, N. R. Deoxyribonucleic and ribonucleic acid homology studies of the genera *Azospirillum* and *Conglomeromonas*. Int. J. Syst. Bacteriol, Washington, v.36, p.80-85, 1986.
- FU, H.; BURRIS, R.H. Ammonium inhibition of nitrogenase activity in *Herbaspirillum seropedicae*. J. Bacteriol, Washington, v.171, p.3168-3175, 1989.
- GRANT, S.G.N.; JESSEE, J.; BLOOM, F.R.; HANAHAN, D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc. Natl. Acad. Sci., USA, Washington, v.97, p.4645-4649, 1990
- ISHIKAWA, J; HOTTA, K. FramePlot: a new implementation of the Frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. FEMS Microbioloy Letter, v.174, Issue 2 , p.251-253, 1999.

- JAGGI, R.; van HEESWIJK, W.C.; WETERHOFF, H.V.; OLLIS, D.L.; VASUDEVAN, S.G. The two opposing activities of adenylyltransferase reside in distinct homologous domains, with intramolecular signal transduction. The EMBO Journal, v.16, p.5562-5571, 1997.
- JIANG, P., NINFA, A.J. Regulation of autophosphorylation of *Escherichia coli* nitrogen regulator II by the PII signal transduction protein. J. Bacteriol., v.181, p.1909-1911, 1999.
- KLASSEN, G.; PEDROSA, F.O.; SOUZA, E.M.; FUNAYAMA, S.; RIGO, L.U. Effect of nitrogen compounds on nitrogenase activity in *Herbaspirillum seropedicae*. Can. J. Microb., Ottawa, v.43, p.887-891, 1997.
- KLASSEN, G.; PEDROSA, F.O.; SOUZA, E.M.; YATES, M.G.; RIGO, L.U. Sequencing and functional analysis of the *nifHnifDnifE* gene cluster of *Herbaspirillum seropedicae*. Fems Microbiol. Letters, v.181, p.165-170, 1999a.
- LIU, M.; TURNER, R.J.; WISTONE, T.R.; SAETRE, A.; DYLLICK-BRENZINGER, M.; JICKLING, G.; TARI, L.W.; WEINER, J.H.; TAYLOR, D.E. *Escherichia coli* TehB Requires S-Adenosylmethionine as a Cofactor To Mediate Tellurite Resistance. Journal of Bacteriology, v.182, n° 22, p.6509–6513, 2000.
- MACHADO, I.M.P.; YATES, M.G.; MACHADO, H.B.; SOUZA, E.M.; PEDROSA, F.O. Cloning and sequence of the nitrogenase structural genes *nifHDK* of *Herbaspirillum seropedicae*. Braz. J. Med. Biol. Res., Ribeirão Preto, v.29, p.1599-1602, 1995.
- MAGASANIK, B. The regulation of nitrogen utilization in enteric bacteria. J. Cell. Biochem., v.51, p.34-40, 1993.
- MALIK, A., NAQVI, S. H. M. & ALEEN, M. I. H. Nitrogen and the environment. Nuclear Institute for Agriculture and Biology, Faisalabad. 1985.



- MERRICK, M.J. Nitrogen control of the *nif* regulon in *Klebsiella pneumoniae*: involvement of the *ntrA* gene and analogies between *ntrC* and *nifA*. EMBO J., Oxford, v.2, p.39-44, 1983.
- MERRICK, M.J.; EDWARDS, R.A. Nitrogen control in bacteria. Microbiol. Rev., Washington, v.59, p.604-622, 1995.
- MUSE, W.B.; BENDER, R.A. Map position of the *glnE* gene from *Escherichia coli*. J. Bacteriol., v.174, p.7876-7877, 1992.
- NIFA, A.J., ATKINSON, M.R., KAMBEROV, E.S., FENG, J., NINFA, E.G. Control of nitrogen assimilation by the NtrB-NtrC two component system of enteric bacteria. *In: Two-component signal transduction*. HOCH, A.; SILHAVY, T.J. (eds.) p.67-88. ASM press, Washington, 1995.
- NOINDORF, Lilian. Clonagem e seqüenciamento parcial do gene para o transportador de amônio em *Herbaspirillum seropedicae*. Curitiba, 55 p. Monografia Setor de Ciências Biológicas – Universidade Federal do Paraná, 2000.
- OLIVER, G., GOSSET, G., SANCHEZ-PESCADOR, R., LOZOYA, E., KU, M.L., FLORES, N., BECERRIL, B., VALLE, F., BOLIVAR, F. Determination of the nucleotide sequence for the glutamate synthase structural genes of *Escherichia coli* K-12. Gene, v.60, p.1-11, 1987.
- PEDROSA, F.O., TEIXEIRA, K.R.S.; MACHADO, I.M.O.; STEFFENS, M.B.R.; KLASSEN, G.; BENELLI, E.M.; FUNAYAMA, S.; RIGO, L.U.; ISHIDA, M.L.; YATES, M.G.; SOUZA, E.M. Structural organization and regulation of the *nif* genes of *Herbaspirillum seropedicae*. Soil Bio.Biochem. v.29. p.843-846, 1997
- PEDROSA, F.O.; BENELLI, E.M.; YATES, M.G.; WASSEM, R.; MONTEIRO, R.A.; KLASSEN, G.; STEFFENS, M.B.; SOUZA, E.M.; CHUBATSU, L.S.; RIGO, L.U. Recent developments in the structural organization and regulation of nitrogen fixation genes in *Herbaspirillum seropedicae*. J Biotechnol, v.91 p.189-95, 2001.

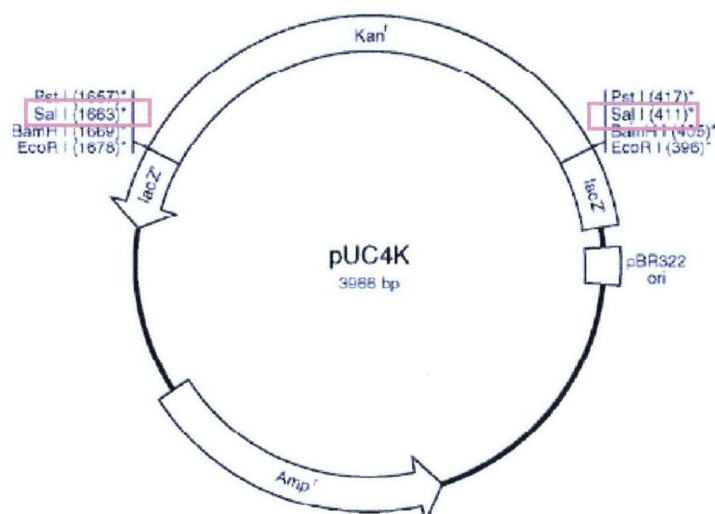
- PERSUHN, D.C., STEFFENS, M.B.R., PEDROSA, F.O., SOUZA, E.M., YATES, M.G., TEIXEIRA, K.R.; RIGO, L.U. The transcriptional activator NtrC controls the expression and activity of glutamine synthetase in *Herbaspirillum seropedicae*. FEMS Microbiology Letters, v.192, p.217-221, 2000.
- PORTER, S.; NORTH, A.K.; KUSTUS, S. Mechanism of transcriptional activation by NtrC. In: Two-component signal transduction. HOCH, A.; SILHAVY, T.J. (eds.) p.147-148. ASM press, Washington, 1995.
- POSTGATE, J.R. Biological nitrogen fixation: Fundamental. Phil. Trans. R. Soc. Lond., Brighton, v.296, p.375-85, 1982.
- RAMOS, Juliana Rocha Lopes. Análises moleculares comparativas de estirpes de *Herbaspirillum* por PFGE, RAPD, RFLP e seqüenciamento do gene 16S rRNA. Curitiba, 2003. 150p. Tese (Doutorado Ciências-Bioquímica) Setor de Ciências Biológicas-Universidade Federal do Paraná.
- REITZER, L.J.; MAGASANIK, B. Isolation of the nitrogen assimilation regulator NR<sub>I</sub>, the product of the *glnG* gene of *Escherichia coli*. Proc. Natl. Acad.Sci. USA., Washington, v.80, p.5554-5558, 1983.
- REITZER, L. J.; MAGASANIK, B. Expression of *glnA* in *Escherichia coli* is regulated at tandem promoters. Proc Natl Acad Sci USA, Washington, v.82, p.1979-1983, 1985.
- REITZER, L.J. Amomonía assimilation and the biosynthesis of glutamine, glutamate, aspartate, asparagine, L-alanine and D-alanine. In: Escherichia coli and Salmonella: cellular and molecular biology. NEIDHARD, F.C. et al. (ed) p.391-407. ASM Press, Washington, 1996.
- SAMBROOK, J.; FRITSCH, E.F. MANIATIS, T. 1989. Molecular cloning: a laboratory manual. 2ed. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press., 1989.
- SANGER, F., NICKLEN, S., COULSON, A. R. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA v.74, p.5463-5467, 1977.

- SCHOLLHORN, R.; BURRIS, R.H. Acetylene as a competitive inhibitor of N<sub>2</sub> fixation. Proc. Natl. Acad. Sci. USA, Washington, v.58, n.213-6, 1967.
- SCHULTZ, J.; COPLEY, R.R.; DOERKS, T.; PONTING, C.P.; BORK, P. SMART: a web-based tool for the study of genetically mobile domains. Nucl. Acids Res., v.28(1), p.231-234, 2000.
- SON, H.S.; RHEE, S.G. Cascade control of *Escherichia coli* glutamine synthetase: purification and properties of PII protein and nucleotide sequence of its structural gene. J. Biol. Chem., v.262, p.8690-8695, 1987.
- SOUZA, E.M. Clonagem, caracterização e sequenciamento dos genes *nifA* e *nifB* de *Herbaspirillum seropedicae*. Curitiba. Tese (Doutorado em Ciências – Bioquímica) – Setor de Ciências Biológicas, Universidade Federal do Paraná, p.245, 1990.
- STADTMAN, E.R.; MURA, E.; CHOCK, P.B.; RHEE, S.G. The interconvertible enzyme cascade that regulates glutamine synthetase activity. *In: Glutamine: Metabolism, Enzymology, and Regulation*. MORA, J. and. PALACIOS, R. (ed) p.41-59. Academic Press, New York, 1980.
- THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucl. Acids Res. v.22, p.4673-4680, 1994.
- VOIGT, Eduardo Luiz. Identificação e Análise Estrutural de Genes a montante do gene *nifA* em *Herbaspirillum seropedicae*. Curitiba, 2000. 74 p. Tese (mestrado Ciências-Bioquímica) Setor de Ciências Biológicas - Universidade Federal do Paraná.
- van HEESWIJK, RABENBERG, M.; WESTERHOFF, H.V.; KAHN, D. The genes of the glutamine synthetase adenylation cascade are not regulated by nitrogen in *Escherichia coli*. Mol. Microbiol., v.9, p.443-457, 1993.

- van HEESWIJK, W.C.; HOVING, S.; MOLENAAR, D.; STEGEMAN, B.; KAHN, D.; WESTERHOFF, H.V. An alternative PII protein in the regulation of glutamine synthetase in *E. coli*. Mol. Microbiol., v.21, p.133-146, 1996.
- WEISS V, MAGASANIK B. Phosphorylation of nitrogen regulator I (NRI) of *Escherichia coli*. Proc. Natl. Acad. Sci.USA, Washington, v.85, p.8919-8923, 1988.
- WEISS, D.S.; BATUT, J.; KLOSE, K.E.; KEENER, J.; KUSTU, S. The phosphorylated form of the enhancer-binding protein NTRC has an ATPase activity that is essential for activation of transcription. Cell, Cambridge, v.67, p.155-167, 1991.
- YAMASHITA, M.M., ALMASSY, R.J., JANSON, C.A., CASCIO, D., EISENBERG, D. Refined atomic model of glutamine synthetase at 3,5Å° resolution. J. Biol. Chem., v.264, p.5780-5787, 1989.

## ANEXOS

### ANEXO 1: Mapa do plasmídeo pUC4K



O plasmídeo pUC4K foi usado como fornecedor do gene de resistência a canamicina. Os sítios *Sal*I utilizados para a restrição estão indicados nos retângulos vermelhos.

Fonte: [www4.amershambiosciences.com](http://www4.amershambiosciences.com)